

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2004, 22:25:40 ; Search time 12355 Seconds
(without alignments)
6083.619 Million cell updates/sec

Title: US-09-522-753-5
Perfect score: 13215
Sequence: 1 MSGSTQLVQATWRATEPRYP.....WDEPKPLLCQYETLSDSE 2517
Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries
Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q=/sgn2_1/USPTO spo01/US09522753/runat_15042004_143737_17439/app_query.fasta_1.2695
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=100 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09522753 @CGN 1 1 10059 @runat_15042004_143737_17439 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_estc1.*
10: gb_estc2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	DB	ID	Description
1	8226	62.2	7372	29	AY412686	AY412686	Homo sapi
2	7109.5	53.8	7013	29	AY412688	AY412688	Mus muscu
3	6533	49.4	6836	29	AY412687	AY412687	Pan trogl
4	2566	19.4	2000	11	BC033087	BC033087	Homo sapi
5	2264	17.1	1362	11	BC020427	BC020427	Homo sapi
6	1643.5	12.4	1966	11	BC026028	BC026028	Homo sapi
7	1634	12.4	1850	11	BC035748	BC035748	Homo sapi
8	1626.5	12.3	1835	11	BC029627	BC029627	Mus muscu
9	1624.5	12.3	1819	11	BC062162	BC062162	Mus muscu
10	1624	12.3	1093	12	BM909096	BM909096	AGENCOURT
11	1603	12.1	1996	11	BC026623	BC026623	Mus muscu
12	1559	11.8	965	12	BM555371	BM555371	AGENCOURT
13	1540	11.7	1126	12	BM553310	BM553310	AGENCOURT
14	1480.5	11.2	1057	12	BM477568	BM477568	AGENCOURT
15	1479.5	11.2	1051	12	BM423558	BM423558	AGENCOURT
16	1471.5	11.1	937	13	BO949384	BO949384	AGENCOURT
17	1448	11.0	846	13	BU557144	BU557144	AGENCOURT
18	1442.5	10.9	984	13	BO690869	BO690869	AGENCOURT
19	1440.5	10.9	1159	12	BM802749	BM802749	AGENCOURT
20	1420	10.7	1022	12	BM910785	BM910785	AGENCOURT
21	1382.5	10.5	1030	12	BM558844	BM558844	AGENCOURT
22	1378.5	10.4	1010	12	BM915686	BM915686	AGENCOURT
23	1374	10.4	887	13	BU542258	BU542258	AGENCOURT
24	1368	10.4	1067	12	BM472005	BM472005	AGENCOURT
25	1366.5	10.3	868	13	BU368971	BU368971	AGENCOURT
26	1348.5	10.2	923	13	BU184403	BU184403	AGENCOURT
27	1346.5	10.2	1038	12	BM910704	BM910704	AGENCOURT
28	1338	10.1	887	14	CA979881	CA979881	AGENCOURT
29	1334	10.1	926	13	BU390462	BU390462	AGENCOURT
30	1332	10.1	875	13	EQ711119	EQ711119	AGENCOURT
31	1331.5	10.1	842	13	BU224569	BU224569	AGENCOURT
32	1326	10.0	1094	12	BM560912	BM560912	AGENCOURT
33	1325.5	10.0	905	13	BO897825	BO897825	AGENCOURT
34	1321	10.0	882	13	BU180236	BU180236	AGENCOURT
35	1312.5	9.9	943	13	BU368972	BU368972	AGENCOURT
36	1305.5	9.9	898	13	BU172348	BU172348	AGENCOURT
37	1305	9.9	1045	12	BM560255	BM560255	AGENCOURT
38	1274	9.6	875	13	BO691710	BO691710	AGENCOURT
39	1262.5	9.6	915	14	CF272442	CF272442	AGENCOURT
40	1260	9.5	1075	12	BM461469	BM461469	AGENCOURT
41	1256.5	9.5	928	13	BU164114	BU164114	AGENCOURT
42	1256	9.5	994	13	BU397973	BU397973	AGENCOURT
43	1253.5	9.5	1066	12	BM471347	BM471347	AGENCOURT
44	1252.5	9.5	797	14	CB723733	CB723733	UI-M-FY0-
45	1250	9.5	1087	12	BG252257	BG252257	UI-M-FY0-
46	1243	9.4	958	12	BG831424	BG831424	602766347
47	1239	9.4	810	12	EG974253	EG974253	602843992
48	1235	9.3	962	13	EQ944732	EQ944732	AGENCOURT
49	1231.5	9.3	1113	12	BM811122	BM811122	AGENCOURT
50	1227.5	9.3	861	13	BU191146	BU191146	AGENCOURT
51	1225.5	9.3	876	12	BG252161	BG252161	602365028
52	1224.5	9.3	763	14	CB248833	CB248833	UI-M-EX0-
53	1216	9.2	845	9	AL884718	AL884718	AL884718
54	1204	9.1	900	13	BQ214358	BQ214358	AGENCOURT
55	1195.5	9.0	1132	12	BM910968	BM910968	AGENCOURT
56	1183.5	9.0	893	13	BU186180	BU186180	AGENCOURT
57	1183	9.0	790	13	BU257626	BU257626	603414849
58	1175	8.9	1032	13	BQ070408	BQ070408	AGENCOURT
59	1161.5	8.8	858	13	EX732682	EX732682	AGENCOURT
60	1158.5	8.8	787	10	BE793487	BE793487	601588814
61	1153.5	8.7	718	14	CA749602	CA749602	UI-M-FY0-
62	1151.5	8.7	704	14	CD355812	CD355812	UI-M-FY0-

63	1124	8.5	804	10	BE728145	601563413	FEATURES	Location/Qualifiers
64	1123	8.5	779	12	BM944466	UI-M-EHOP		
65	1116	8.4	811	13	BX771375	BX771375	source	1..7372 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>7372 /gene="NCOR2" /locus_tag="HCM4640"
66	1114	8.4	774	14	CB521565	UI-M-GH0-		
67	1111	8.4	651	12	BM783748	K-EST0061	gene	
68	1110.5	8.4	683	14	CF540568	UI-M-GV0-		
69	1107	8.4	622	9	AI830862	WJ61509.X	ORIGIN	
70	1105	8.4	657	10	BF530324	602071630		
71	1104.5	8.4	902	12	EG485444	602505328	Alignment Scores:	Length: 7372 Matches: 1708 Conservative: 1 Mismatches: 748 Indels: 0 Gaps: 29
72	1099.5	8.3	825	13	CB216442	NISC nq05		
73	1098.5	8.3	825	13	BX732681	BX732681	Pred. No.: 0	
74	1090	8.2	922	10	BE542336	601067079		
75	1087	8.2	682	13	BX869832	BX869832	Score: 8226.00	
76	1085	8.2	790	12	BI081252	602879279		
77	1081	8.2	729	13	CA325207	UI-M-FY0-	Percent Similarity: 69.56%	
78	1074	8.1	809	13	BU475754	603471578		
79	1073.5	8.1	788	12	EM011980	603636389	Best Local Similarity: 69.52%	
80	1070.5	8.1	688	10	BE391091	601286040		
81	1070	8.1	664	10	BF058781	60112484	Query Match: 62.25%	
82	1062.5	8.0	1088	13	BU108220	603356210		
83	1053	8.0	592	10	BF436351	7p06e11.X	DB: 29	
84	1039.5	7.9	820	12	BI735071	603356210		
85	1038	7.9	611	10	AW438580	x882e10.X	US-09-522-753-5 (1-2517) x AY412686 (1-7372)	
86	1038	7.9	624	12	BG391632	602417893		
87	1038	7.9	668	14	CB518985	UI-M-GH0-	QY	1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro
88	1037	7.8	925	14	CD244013	AGENCOURT		
89	1035	7.8	1030	12	BG119261	602349134	Db	1 ATGTCGGGATCCACACAGCTGTGGCAGAGCTGGAGGGCCACTGAGCCCGCTACCCG
90	1033.5	7.8	654	10	AW701437	uq86b05.Y		
91	1032	7.8	725	14	CA317171	UI-M-FW0-	QY	21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu
92	1030	7.8	614	12	BM771666	K-EST0055		
93	1027.5	7.8	810	10	BF606925	MY2 00011	Db	61 CCCCACAGCTTTCCTACCCAGTGCAGATCGCCCGACGACACGAGCTCGGGCTCCTG
94	1025	7.8	612	9	AI523558	th08C09.X		
95	1020.5	7.7	898	10	BF339827	602038834	QY	41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln
96	1017.5	7.7	637	12	BI428444	fr49a07.Y		
97	1017.5	7.7	760	13	BU143568	603229009	Db	181 CCCCAGCGCGAGGCGCTCCCTGCTGTCTGAGTTCAGCCGGGAATGAACGNNNNNN
98	1014	7.7	612	10	AW674918	bb30f04.Y		
99	1012	7.7	610	10	BE675456	7f09d11.X	QY	81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet
100	1011.5	7.7	676	13	BY763864	BY763864		
ALIGNMENTS								
7372 bp DNA linear GSS 16-DEC-2003								
RESULT 1	AY412686	Homo sapiens NCOR2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.						
LOCUS	AY412686							
DEFINITION	AY412686.1							
ACCESSION	AY412686							
VERSION	AY412686.1							
KEYWORDS	GSS.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	1 (bases 1 to 7372)							
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.							
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios							
JOURNAL	Science 302 (5652), 1960-1963 (2003)							
PUBMED	14671302							
REFERENCE	2 (bases 1 to 7372)							
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.							
TITLE	Direct Submission							
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA							
COMMENT	This sequence was made by sequencing genomic exons and ordering							

Db 781 TCCGACCCCGCAGTATCATGAGAACATCAAAATAAACCGAGCGCATCGGAAGAAGCTA 840
Qy 281 IleLeuTyrPheLysArgAsnHisAlaArgLysGlnTyrLysGlnLysPheCysGln 300
Db 841 ATCTTGTAATTCACAGAGAGGAAATCACCTCGAAACAATGGAGCAGAAGTTCTGCCAG 900
Qy 301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
Db 901 CGCTATGACAGCTCATGAGGCGCTGGGAGAGAGAGGTTGGAGCGCATCGAGAACAAACCCC 960
Qy 321 ArgArgArgAlaLysGlnSerLysValArgGlnTyrTyrGluLysGlnPheProGluIle 340
Db 961 CGCGCGCGGCGCAAGGAGAGCAAGTGGCGAGTACTACGAGAACAGCTTCCCTGAGATC 1020
Qy 341 ArgLysGlnArgGluLeuGlnArgMetGlnSerArgValGlnArgLysSerGly 360
Db 1021 CGCAAGCAGCGGAGCTCAGAGCGCATGCAAGCAGGTTGGCCAGCGGGCAGTGGG 1080
Qy 361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
Db 1081 CTGTCCATGTCCGCGCGCCGAGCAGCAGAGGTGTCAAGATCATCGATGGCTCTCA 1140
Qy 381 GlnGlnGlnAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
Db 1141 GAGCAGGAGAACCTGGAGAGCAGATGCCAGCTGGCGGTGATCCCGCCCATGCTGTAC 1200
Qy 401 AspAlaAspGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAsnProMet 420
Db 1201 GACGCTGACACAGCGGATCAAGTTTCATCAATGAAACGGGCTTATGGCCAGCCCCATG 1260
Qy 421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
Db 1261 AAGTGTACAAAGACCGCCAGTCAATGAATGTGAGTGTGAGCAGAGAGAGACCTTC 1320
Qy 441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
Db 1321 CGGAGAGAAANNN 1380
Qy 461 LysThrValAlaGluCysValLeuTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
Db 1381 NNNACAGTGGCTGAGTGGCTCTCTATTACTACCTGACTAAGAGAAATGAGAACTATAAG 1440
Qy 481 SerLeuValArgArgSerTyrArgArgGlyLysSerGlnGlnGlnGlnGlnGln 500
Db 1441 AGCTGTGTGACGCGAGCTATCGCGCGCGCGAGACGACGAGNNNNNNNNNNNNNNNN 1500
Qy 501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluLys 520
Db 1501 NNN 1560
Qy 521 AspGluLysGluLysGluAlaGluLysGluGluLysProGluValGluAsn 540
Db 1561 NNN 1620
Qy 541 AspLysGluAspLeuLeuLysGluLysThrAspThrSerGlyGluAspAsnAspGlu 560
Db 1621 NNN 1680
Qy 561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
Db 1681 AAGGAGGCTGTGGCTCCAAAGGCGCAAAACTGCCAACAGCAGCGGGAAGACGCAAGGC 1740
Qy 581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluAlaIleThrProGlnGln 600
Db 1741 CGCATCACCTCAATGGCTAATGAGGCCAACAGCAGGAGGCGCATCACCCCGCAGCAG 1800
Qy 601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTyrThrGluGluMet 620
Db 1801 AGCGCCGAGCTGGCTTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAGAGAAATG 1860
Qy 621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640

Db 1861 GAAACAGCAAGAAGGTCTCTCTGGAAACACGCGCGCAACTGGTCGGCCATCGCCCGGATG 1920
Qy 641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
Db 1921 GTGGCTCCAAGACTGTGTCCAGCTGTGAAGAACTTCTCACTCACTACAGAGAGGAG 1980
Qy 661 AsnLeuAspGluIleLeuGlnGlnHisLysLysLeuLysMetGluLysGluArgAsnAlaArg 680
Db 1981 AACCTCGATGAGTCTTGCAGCAGCACAAAGCTGAAGATGGAGAGAGAGAACCGCGG 2040
Qy 681 ArgLysLysLysLysAlaProAlaAlaLaserGluGluAlaAlaPheProProValVal 700
Db 2041 AGGAAGAAGAAGAGCGCGCGCGCAGCAGGAGGCTGCATTCCCGCCCTGTGTG 2100
Qy 701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluMetValGluGlu 720
Db 2101 CAGATGAGGAGATGGAGCGCTCGGGCTGAGCGGAATGAGGAGAGATGTTGGAGGAG 2160
Qy 721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
Db 2161 GCTGAAGNN 2220
Qy 741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaLys 760
Db 2221 NNN 2280
Qy 761 AspThrGlyGlnAsnGlyProLysProAlaThrLeuGlyAlaAspGlyProProPro 780
Db 2281 NNN 2340
Qy 781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
Db 2341 NNN 2400
Qy 801 GluAlaThrGlyAlaProThrProProAlaProProSerProSerAlaProProPro 820
Db 2401 NNN 2460
Qy 821 ValValProLysGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
Db 2461 NNN 2520
Qy 841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
Db 2521 GAGCAGCAGAGAGCCCGCGCTGAGGAGCTGGCAGTGGACACAGGGAAGCGCCGAGGAG 2580
Qy 861 ProValLysSerGluCysThrGluGluAlaGluGluGluProAlaLysGlyLysAla 880
Db 2581 CCCGTCAAGAGCGAGTGCAGCAGAGAGCGGAGGCGCGCGCAAGGGAAGGACGCG 2640
Qy 881 GluAlaAlaGluAlaThrAlaGluGluLysAlaGluLysAlaGluLysGlySer 900
Db 2641 GAGCGCTGTAGGCCACCGCGCGAGGGCGCTCAAGGCAGAGAGAGAGGCGGCGGAGC 2700
Qy 901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
Db 2701 GGCAGGCGCCACACAGCAAGAGCTCGCGCGCGCGCCCGCAGGACAGCGACTCCAGTGTACC 2760
Qy 921 CysSerAlaAspGluValAspGluAlaGluGlyAspLysAsnArgLeuLeuSerPro 940
Db 2761 TGCAGTCAGACGAGGTGGATGAGCGCGAGGCGCGCGCAAGAGACCGGCTGTGTGCCA 2820
Qy 941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
Db 2821 AGGCCAGCTCTTCAACCCGACTGGCGACCCCGGCGCAATGCTCTACCCAGAGGCA 2880
Qy 961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaIleProProIleGlnValThrLys 980
Db 2881 CTGGACCTGAAGCAGCTGAAGCAGCAGCGCTGCATCCCCCCCCCATCCAGTCAACAAA 2940
Qy 981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
Db 2941 GTCCATGAGCCCCCGGAGGAGCGAGCTCCCAAGCAGCTCCCGCCAGCCCGCCCGC 3000

Db 5161 GCACTCAACTAGCTGCGGGTCCCGAGGACATCATCGACCTGTCCAAAGTGCCACACCTG 5220
Qy 1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
Db 5221 CCGTGTCTGTCGCCCCGACACCGAGCACCCGAGCACCGCATGGACCGCTTGCCTAC 5280
Qy 1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
Db 5281 CTCGCCACCGCGCCCGACCTTCAGCAGCGCCACACAGCAGCTCCCACTCTCCCCAGGA 5340
Qy 1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
Db 5341 GGTCCACACACTTGACAAAAACCAACCAACGCTCTCGTCGAGCGGGAGCGAGACCGG 5400
Qy 1801 AspArgGluArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
Db 5401 GATCAGAGCGGGACCGGGATCGGAGCGGGAGGAGGAGTCCATCTCCTCACCACGACG 5460
Qy 1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
Db 5461 GTGAGCAGCAGCACCCATCTGAGACCTGTGATACAGAGCAGAGCAGCGGCGAGCGGACG 5520
Qy 1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
Db 5521 AGCGCGGGGTGGGGGAGCAGCAGCGCGCGCTCCCATCTCCCATGCGCCACCCAGCAC 5580
Qy 1861 SerProIleSerProArgThrGlnAspAlaLeuGlnArgGlnArgProSerValLeuHisAsn 1880
Db 5581 TCGCCCATCTCCCTCGGACCCAGATGCCCTCCAGCAGAGACCCAGGTGCTTCACAAC 5640
Qy 1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
Db 5641 ACAGGCATGAAGGTATCATCACCGCTGTGGAGCCACAGCCCGCTCTGAGGTCC 5700
Qy 1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
Db 5701 ACCTCCACTCTCACCGCTTCGCGCGCTGCCACATTCACCACTGCCACCACTGCCCA 5760
Qy 1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
Db 5761 CTGGGGGACCCCTCGATGGGTCTACCTTACCTCATGGAGCCCGCTTCTGCTGCCAAG 5820
Qy 1941 GluAlaProArgValAlaArgProGluArgProAlaAlaAspThrGlyHisAlaPheLeu 1960
Db 5821 GAGGCCCGCGGTGCGCCGCGAGCGCGCCGAGCAGACACCGGCTTCCTTCCTC 5880
Qy 1961 AlalysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
Db 5881 GCCAAGCCCGCAGCGCGCTCGGGCTGGAGCCCGCTCTCCCGCAGCAAGGGCTCGGAG 5940
Qy 1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaAtqThrProAlalys 2000
Db 5941 CCGCGCGCCCTAGTGCTCTGTCTGGCCAGCCACCATCGCGCCGACCCCTCGCGAAG 6000
Qy 2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
Db 6001 AACCTCGACCTCACCACCGCAGCGCGGACCGCGCGGCGCCACCTGCTCGGCTCGGAC 6060
Qy 2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnLeuLeuLeuArgSer 2040
Db 6061 CCGCACCGGAAAAAGACTCAAAAGTAAACCTTTTCATCCAGGAACCTGGAACCTCGTTCT 6120
Qy 2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
Db 6121 CTGNN 6180
Qy 2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluLeuLeuAspLysSer 2080
Db 6181 NNN 6240
Qy 2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
|||||

Db 6241 NNN 6300
Qy 2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
Db 6301 GCCACCTCCACACCTGCGGCGCTGCTGAGAGCAGCCCTGCTCCAGCCCGTGTCTC 6360
Qy 2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
Db 6361 CAGACCGCCCGAGGGGTCAAAGGTCAACAGGGTGTGTCACCTCGCCGACGACATCACT 6420
Qy 2141 GluValIleThrGlnAspTyrThrArgHisIleProGlnGlnLeuSerAlaProLeuPro 2160
Db 6421 NNN 6480
Qy 2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
Db 6481 NNN 6540
Qy 2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
Db 6541 NNN 6600
Qy 2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGluAsp 2220
Db 6601 NNN 6660
Qy 2221 GlyIleGluProValSerProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
Db 6661 GGTATTGACCTGTGTCGCCACCGAGGGCATGACGGAGCCAGGCACTCCCGGAGTGTCT 6720
Qy 2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
Db 6721 GTGTACCGCTGTGTACCGGGATGGGAAACAGACGGAGCCAGCAGGATGGGTCCAAG 6780
Qy 2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
Db 6781 TCTCCAGCAACACCGAGCCCGCCAGCTTCTTTCAGCAAGCTACCGAGAGCAACTCC 6840
Qy 2281 AlaMetValLysSerLysLysGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
Db 6841 GCCATGTCAAGTCCAGAGCAAGAGATCAACAAGAAGCTGAACACCAACCCGGAAT 6900
Qy 2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
Db 6901 GAGCTGAATACANN 6960
Qy 2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
Db 6961 NNN 7020
Qy 2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
Db 7021 NNN 7080
Qy 2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
Db 7081 NNN 7140
Qy 2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
Db 7141 NNN 7200
Qy 2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
Db 7201 GGGAGGCGCAAGGTCTCTGGCAGACCCAGCAGCGAAAGCAAGTCCCGGCCCGGGC 7260
Qy 2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
Db 7261 CTGCACTCGGGACCGGCCACCTCTGCTCTCTCCTCAGTGCACTCGAGGGAGACTGCAAC 7320
Qy 2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAla 2457
Db 7321 CGCCGAGCGCGCTCACCAACCGCGGTGTGGAGGAGCAGGCCCTCTGTCGCA 7371
|||||

```
RESULT 2
LOCUS AY412688 7013 bp DNA linear GSS 16-DEC-2003
DEFINITION Mus musculus NCOR2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY412688
VERSION AY412688.1 GI:39768653
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 7013)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 7013)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..7013
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>7013
/gene="NCOR2"
/locus_tag="HCM4640"
gene
ORIGIN
Alignment Scores: 3,86e-284 Length: 7013
Pred. No.: 7109.50 Matches: 1509
Score: 63.77% Conservatives: 59
Percent Similarity: 61.37% Mismatches: 768
Best Local Similarity: 53.80% Indels: 125
Query Match: 29 Gaps: 17
DB:
US-09-522-753-5 (1-2517) x AY412688 (1-7013)
QY 1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
Db 1 ATGTGAGGATCCACACAGCCTGTGGCAGACATGCGCGGCTGCTGAGCCCGCTACCCA 60
QY 21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
Db 61 CCCCATGGCATCTCTTACCCCGGTGAGATAGCCCGTCCACACGACGTGGGGTGCTT 120
QY 41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleLeuGln 60
Db 121 GAGTACCAACACACACCCCGTGACTACCTCACACCTGTCAACCCGTTCCATCATCCAG 180
QY 61 ProGlnArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
Db 181 CCACAGAGGAGCGGCCCTCACTGCTGTCAGATTCAGCTCGGAGTGACGNNNNNN 240
QY 81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 299
QY 101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
```

```
300 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 359
121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
Db 360 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 419
141 LeuThrGlyLysLeuGluProValSerProProSerProHisThrAspProGluLeu 160
Db 420 CTGGCAGGCAAGCTGGAGCCTGTGTACCTCCAGTCCCGCACGCTGACCTGAGCTA 479
161 GluLeuValProProArgLeuSerLysGluLeuIleGlnAsnMetAspArgValAsp 180
Db 480 GAGCTGGCGCCATCTCGACTGTCCAAGGAGGAGTGTCCAGAACATGGACCGCTGGAC 539
181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnLeu 200
Db 540 CGTGAGATCACCATGTTAGACGACGATCTCCAGCTGCTGAGAGAGACGACAGCTTG 599
201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
Db 600 GAGGAGGAGCGCCCAAGCCGCCGAAACCGAGAGAGCTGTGTCCGCCACCCATAGAA 659
221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
Db 660 TCAAGACGACCGAAGCCTGGTCCAGATCATCTACGATGAGAACCGNNNNNNNNNNNN 719
241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
Db 720 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 779
261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
Db 780 TCTGACACACCGCAGTACCATCAAAACATCAAAATAAACACGCGCATGCGGAGAGCTG 839
281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
Db 840 ATCTTGTACTTTAAGCGGAGGAAACCAACGCGCGCAGCAGTGGGAAACAGCGCTTCTGCCAG 899
301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
Db 900 CGCTATGACCAGCTCATGGAGCGCTGGGAGAGAGAGTGTAGCGCATAGAGAACATCCG 959
321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
Db 960 CGAAGAGGCGCCAAAGGAGAGCAAGGTGAGGAGTACTACGAGAAACAGTTCCTCCGAGATC 1019
341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
Db 1020 CGCAAGCAGCGGNGCTGTCAGGAGCGCATGSCAGAGCGGTGGCGCAGCGTGGCAGTGGG 1079
361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
Db 1080 CTCTCCATGTCGGCTGCCCGCAGTGTGAGTGTCTGAGATCATGTGATGTGCTGTCT 1139
381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
Db 1140 GAGCAGAGAAACCTGGAGAGACAGATGCGCAGCTGCCCGTGTATCCCGCCCATGTTGTAC 1199
401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
Db 1200 GACGCGGACGACGAGAGATCAAGTTCATCAACATGAATGAGACTCATGTGATGACCCCATG 1259
421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
Db 1260 AAGGTCTCAAGAGCCGTCAGTTACCAACATGTGGAGCGGAGCAGGAGGAGGACACCTTC 1319
441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
Db 1320 CGTGAGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1379
461 LysThrValAlaGluCysValLeuTyrTyrTrpLeuThrLysLysAsnGluAsnTyrLys 480
```

Db 1380 NNNACGGTCGTGAGTGTCTCTATTACTACCTGACCAAGAAATGAATAATTACAAG 1439
Qy SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGln 500
Db AGCTTGGTGGCGGAGCTATCGCGCGCTGCGAAGAGCCAGNNNNNNNNNNNNNNNN 1499
Qy GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 520
Db NNN 1559
Qy AspGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 540
Db NNN 1619
Qy AspLysGluAspLeuLeuLysGluLysThrAspThrSerGlyLysAspAspGlu 560
Db NNNNNNNNNNNNNNNNNNNNGAAGAGACAGACACACTTCTGGCGAGGACACGATGAG 1679
Qy LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgLysGly 580
Db AAAGAGGCGGTGGCTCCAAAAGCGCGCAAACTGCCAACAGCAAGGCGCGCAAGGC 1739
Qy ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluAlaIleThrProGlnGln 600
Db CGTATCACGCGCTCCATGGCCAACAGGCGCAACCATGAGGAGACAGCCACCCACAGCAA 1799
Qy SerAlaGluLeuAlaSerMetGluLeuAsnGlnSerSerArgTyrThrGluGluMet 620
Db AGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGTGGACTGAGGAGAGATG 1859
Qy GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
Db GAGACAGCAAGAAGGCGCTCTGGAAACATGGGAGGAATGGTCAAGCATTTGCCGCGCATG 1919
Qy ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysArgGln 660
Db GTGGCTCCAGACCGTGTCCAGGTGAAGAACTTCTACTTCAACTCAAGAAGAGGCGAG 1979
Qy AsnLeuAspGluIleLeuGlnGlnHisLysLysLeuLysMetGluLysGluArgAsnAlaArg 680
Db AACCTGGACGAAATCTTCAGCAGCACCAAGCTAAGATGGAGAGGAGGAGNACGCTCGG 2039
Qy ArgLysLysLysLysAlaProAlaAlaSerGluGluAlaAlaPheProValVal 700
Db AGGAAGAAGAAGAACCCAGCTCGCGGAGGAGGAGACAGCTTCCACCTGCCGCT 2099
Qy GluAspGluMetGluAlaSerGlyValSerGlyAsnGluGluMetValGluGlu 720
Db GAGGACGAAGAGATGGAGCATCAGCGCAAGTGCATGAGGAAGAGCTGGCGGAGGAG 2159
Qy AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
Db GCAGAGNN 2199
Qy ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
Db NNN 2249
Qy AspThrGlyGlnAsnGlyProLysProAlaThrLeuGlyAlaAspGlyProProPro 780
Db GAG----- 2252
Qy GlyProProThrProProArgThrSerArgAlaProIleGluProThrProAlaSer 800
Db ----- 2252
Qy GluAlaThrGlyAlaProThrProProAlaProProSerProSerAlaProProPro 820
Db ----- 2252
Qy ValValProLysGluGluLysGluGluThrAlaAlaAlaProProValGluGluGly 840
Db -----CAGAAAGTCTGAGCGCGAGGAG-----ATCGATGTGGGA 2285

Qy 841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
Db AAGCCAGGAGGAGCGGAGCGCTCTGAGGAG-----CCCCCGGAG 2324
Qy ProValLysSerGluCysThrGluGluAlaGluGluGlyProAla-----LysGlyLysAsp 879
Db AGTGTAAAGAGTGCACCAAGAGGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAAGGC 2384
Qy AlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGly 899
Db ACAGAGGCCATTGAACCTGTCTGAGGCCACCTTAAGGTGGAG-----GAGCTGGT 2438
Qy SerGlyArgAlaThrThr-----AlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918
Db AGCAAGGCAGCTGTGACCAAGGGTTCAGCTCAGTGCACCCAGGACAGTACTCCAGT 2498
Qy AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938
Db GCCACCTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 2558
Qy SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
Db TCACCAAGGCCAGCTCTCCTCAGCGCTGGAGATCCCGGGCCAGTACTCTGCCCCAG 2618
Qy LysProLeuAspLeuLysGlnLysGlnArgAlaAlaIleProProIleGlnVal 978
Db AAGCGCTGGACCTGAAGCAGCTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 2678
Qy ThrLysValHisGluProProArgGluAspAlaProThrLysProAlaProProAla 998
Db ACCAAGTCTCATGAGCCCCCGGAGGACACAGTATCCCCCAAGCCAGTATCCCCCTGTG 2738
Qy ProProProProGlnAsnLeuGlnProGlnSerAspAlaProGlnGlnProGlySerSer 1018
Db CTTCCACCGCAGCAGCTTACAGCAGAGGCTGACGTCTCAGCAGTCCGAGGAAAGT 2798
Qy ProArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaGlu 1038
Db CCACGTGGCAAGTCCCGCAGCCAGTCTCTCCGAGAAAGAGGNNNNNNNNNNNNNN 2858
Qy AlaGlnLysLeuProGlyAspProCysTrpThrSerGlyLeuProPheProValPro 1058
Db NNN 2918
Qy ProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAla 1078
Db NNN 2978
Qy ProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg 1098
Db NNNNNNGTCACCGCTGCTCTGGGCTTCCACGATAGTGCAGCGCGCTCTGCGCACGT 3038
Qy ProProThrIleSerAsnProProLeuLeuSerSerAlaLysHisProSerValLeu 1118
Db CCCCC---AFTCTAAACCCCCACCTCTCTCTCTGCAAGCATCCCGGCGTACTT 3095
Qy GluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSer 1138
Db GAGAGGACGTGGGTGGCATCTCCAGGGATGTCAGTCCAGCTTCTGTGCTCCTCACTCA 3155
Qy GluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspPro 1158
Db GAGCATGCCAAGGCCCATGGGCTCTCACCATGGGCTGCCCTTGGCGTGACCT 3215
Qy LysLysLeuAlaProPheSerGlyValLysGlnGlnLeuSerProArgGlyGlnAla 1178
Db AAGAAGCTGGCCCCCTTTCAGTGGAGTGAAGCAGAAACAGTGTGCTCCCTCGGGGTGAGCT 3275
Qy GlyProProGluSerLeuGlyValProThrAlaGlnAlaSerValLeuArgGlyThr 1198
Db GGGCCACTGTAAGTGTGGGAGTCCCACTGCTCAGGAGACCTCTGTGCTGAGAGGGACA 3335

QY 1199 AlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValPro 1218
 DB 3336 GCATGGGCTCCGCCACAGTGGAGCATCACAGGGCTCCAGTAGTACCGGGCTGCA 3395
 QY 1219 SerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeu 1238
 DB 3396 GACGGCCCC-----AGCTACAGAGGCTCTATCACCGACGCGCGCCGACAGCTCCTC 3449
 QY 1239 TyrLysGlyThrIleThrArgIleGlyGluAspSerProSerArgLeuAspArgGly 1258
 DB 3450 TACAGGGTACCATCAGCAGATCGTGGTGAGACAGCCCAAGTCCCTTGACCGGCA 3509
 QY 1259 ArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeu 1278
 DB 3510 CAGAGGACACCTCCGCCAGGGCCATGCTCATCTAGGGCAAGAAAGGCCACGTCCTA 3569
 QY 1279 SerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSer 1298
 DB 3570 TCCTATGAAGNN 3629
 QY 1299 GlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgVal 1318
 DB 3630 NNN 3689
 QY 1319 GlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGlu 1338
 DB 3690 NNN 3746
 QY 1339 ArgHisSerProHisHisLysLysGluHisHisIleArgGlySerIleThrGlnGly 1358
 DB 3747 CAGCAGAGCCCC---CACCTCAAGAGAGCATCATCATCCGAGGCTCCATCAGCAAGN 3803
 QY 1359 IleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeu 1378
 DB 3804 NNNNN---NN 3862
 QY 1379 LysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLys 1398
 DB 3863 NNNNNN-----NN 3895
 QY 1399 ThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrVal 1418
 DB 3896 NNN 3955
 QY 1419 LysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGlu 1438
 DB 3956 NNN 4015
 QY 1439 LeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLys 1458
 DB 4016 NNN 4075
 QY 1459 TyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerIleGly 1478
 DB 4076 NNN 4135
 QY 1479 SerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAla 1498
 DB 4136 AGCCCCGGCGGCTTCCCTGCGCTGACCCCGCTGACATATGGTGTGATGCCGGCA 4195
 QY 1499 LeuGluArgAlaCysTyrGluGluSerLysSerArgProGlyThrAlaSerSer 1518
 DB 4196 CTGGAGCGTGTCTGCTATGAAGAGAGTCTGAAGAGCGGTCAGGACAGCAGTGGTGCA 4255
 QY 1519 GlyGlySerIleAlaArgAlaProValIleValProGluLeuGlyLysProArgGln 1538
 DB 4256 GGGGGCTCCATCACACGTGGGGCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4315
 QY 1539 SerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySer 1558
 DB 4316 AGCCCCAGCTTACGAGAGACCGGGGACCCCTTCCAGTCACTCCATCCAGCTGGCTCC 4375
 QY 1559 ProValThrMetArgGluProThrProArgLeuGlnGluGlySerSerSerLys 1578

DB 4376 CCTGTGACCAGGAGGACCCAGCCAGCGCTTCAGGAAGNNNNNNNNNNNNNNNNNN 4435
 QY 1579 AlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSer 1598
 DB 4436 NNN 4495
 QY 1599 ThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyVal 1618
 DB 4496 NNN 4555
 QY 1619 SerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIlePro 1638
 DB 4556 NNN 4615
 QY 1639 ArgGlyIleProLeuAspAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsn 1658
 DB 4616 NNN 4675
 QY 1659 ProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAla 1678
 DB 4676 CCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGTACCTTGACACGCGGCC 4735
 QY 1679 LeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisAsn 1698
 DB 4736 CTGGAGAACCGGCAGACCATCATCAATGACTACATCATCTCGACAGATGACCCACAC 4795
 QY 1699 ThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSer 1718
 DB 4796 GCTGCTCCGCCATGGCCAGCGTGTGTGACATGCTGAGGGTCTGTCCACCGGAGAGTCC 4855
 QY 1719 SerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValPro 1738
 DB 4856 TCGCTGGCCCTCAATTATGCGCTGGCCCAAGAGGCAATTATCGACCTGTCTCCCAAGTGCCA 4915
 QY 1739 HisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeu 1758
 DB 4916 CACCTGCGCGTGTGTGTGTCACCAAGCCAGGACCCCTGCCACGCCCATCGACGCGCTT 4975
 QY 1759 AlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerProLeuSer 1778
 DB 4976 GCCTACCTCCCATCTGCGCCCGCCACCGCTTCAGCAGCGCCACAGTAGTAGTCCACGCTGTCC 5035
 QY 1779 ProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerGluArgGluArg 1798
 DB 5036 CCAGAGGCGCCCATCTACCTAGTAAACCACTATCCACATCTTCATCGGAGCGGGAACGG 5095
 QY 1799 AspArgAspArgGluArgAspArgGluArgGluLysSerIleLeuThrSerThr 1818
 DB 5096 GAACTGAGCGGGAACGAGAC-----AAGTCCATCTCTCAGCTCTTACC 5137
 QY 1819 ThrThrValGluHisAlaProIleTyrArgProGlyThrGluGlnSerSerGlySerSer 1838
 DB 5138 ACTACGTGGAGCATGACCCCATCTGGAGACCTGTGACGAGCAGCAGCAGCGGGCT--- 5194
 QY 1839 GlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHis 1858
 DB 5195 -----GGGGGAGCAGCGCGCCCGCTCCACACCC-----CAC 5227
 QY 1859 GlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnArgProSerValLeu 1878
 DB 5228 CAGCACTCGCCATCTCCCGCCGACCCAGGACCGCTTGCAGCAGAGAGCCCGCTGCTG 5287
 QY 1879 HisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeu 1898
 DB 5288 CACAACACGAGCATGAGGGGTGTGCTACCTCCGTGGAAACCCCGCAGCCCGCTGCTG 5347
 QY 1899 ArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHis 1918
 DB 5348 AGGTCCACCTCCACCTCTGCGCTGCTGCGCCAGTGCACATTCCACCTGCCACCCAC 5407
 QY 1919 CysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeu 1938

5408	Db	TGCCCACTTGTGTGGCACCCCTTGAAGGGGTCTACCCCTCATCTGAGAGCCCGTCTCTCTTTA	5461
1939	Qy	ProLysGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAla	1958
5468	Db	CCCAAGGAGACCTCTCGGGTCGCCGCCGAGCGCCCCGTGTGGACGCTGGCCATGCC	5527
1959	Qy	PhelLeuAlaLysProProAlaAArgSerGlyLeuGluProAlaSerSerProSerLysGly	1978
5528	Db	TTCTCTACCAACCCCGGCCGCGG-----GAGCCCGCTCTCACCAGCAAGAGC	5578
1979	Qy	SerGluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrPro	1998
5579	Db	TCCGAGCCCGATCCTTAGCACCCCGAGCTCCAGCCACAGCCATCGCCCGCACCCCA	5638
1999	Qy	AlaLysAsnLeuAlaProHisHisAlaSerProAspProProAlaProAlaSerAla	2018
5639	Db	GCAAGAGCCTTGACCCCACTGCCAGTCCGAGCCCGCGCGG---CCCACCTCGGCC	5695
2019	Qy	SerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeu	2038
5696	Db	TCAGATCTGCACCCGAGAAAGACTCAAAGTAAACCTTTTCATCCAGGAATGGAACTC	5755
2039	Qy	ArgSerLeuGlyTyThrHisGlySerSerTySerProGluGlyValGluProValSerPro	2058
5756	Db	CGTTCTCTGGNN	5815
2059	Qy	ValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAsp	2078
5816	Db	NN	5875
2079	Qy	LysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGly	2098
5876	Db	NN	5935
2099	Qy	GluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerPro	2118
5936	Db	GAGCTCTCCCATCTCCACATCTCGGCCACTGCCCAGAGCCAGCCCTCATCCAGCCCA	5995
2119	Qy	LeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHis	2138
5996	Db	CTCTCCAGACTGCCCAGGCATCAAAGTCAACAGAGGTGTGTCAACCTTGCTCAGCAC	6055
2139	Qy	IleSerGluValIleThrGlnAspTyThrArgHisHisProGlnGlnLeuSerAlaPro	2158
6056	Db	ATCAGCNN	6115
2159	Qy	LeuProAlaProLeuTySerPheProGlyAlaSerCysProValLeuAspLeuArgArg	2178
6116	Db	NN	6175
2179	Qy	ProProSerAspLeuTyLeuProProAspHisGlyAlaProAlaArgGlySerPro	2198
6176	Db	NN	6235
2199	Qy	HisSerGluGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGly	2218
6236	Db	NN	6295
2219	Qy	GluAspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArg	2238
6296	Db	GAGGATCCCATTTGAGCCTGTGTCCCAACAGAGGGCATGACTGAGCCAGACATGCTCGG	6355
2239	Qy	SerAlaValTyProLeuLeuTyArgAspGlyGluGlnThrGluProSerArgMetGly	2258
6356	Db	AGCCCTGTGTACCCACTTGCTGTATCGAGACGGGGAACAGGGCGAGCCCAACAGATGGGC	6415
2259	Qy	SerLysSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSer	2278
6416	Db	TCTAAGTCTCCAGGCAACACCGACCGCCCGAGCCTTCTTCAGTAAGTCACTGAGAGC	6475
2279	Qy	AsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsn	2298
6476	Db	AACTCCGCCATGGTGAAGTCCAAAGAACGAGAGTCAAAGAAATCAAACCCCAAC	6535

[illegible]

Mon Apr 19 08:15:09 2004

Alignment Scores:		2.78e-260	Length:	6836
Pred. No.:	6533.00	Matches:	1407	
Score:	61.81%	Conservative:	1	
Percent Similarity:	61.81%	Mismatches:	870	
Best Local Similarity:	49.44%	Indels:	0	
Query Match:	29	Gaps:	0	
DB:				
US-09-522-753-5 (1-2517) x AY412687 (1-6836)				
QY	27	ProValGlnIleAlaAArgThrHisThrAspValGlyLeuLeuGluTyrGlnHisHisSer	46	
DB	2	CCAGTGCAGATTGCCCGGACGACACGAGCTCCGGCTCTCGAGTACACGACCACTCC	61	
QY	47	ArgAspTyrAlaSerHisLeuSerProGlySerIleleGlnProGlnA-rgArgPro	66	
DB	62	CGCGACTATGCTCCCACTGTGCGCCGCTCCATCATCCAGCCACCGCGGAGGCC	121	
QY	67	SerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGlnGluLeuHisLeuArgPro	86	
DB	122	TCCCTGCTGTCTGAGTTCAGCCCGGGAATGAACGNNNNNNNNNNNNNNNNNNNNNN	181	
QY	87	GluSerHisSerTyrLeuProGluLeuGlyLysSerGluMetGluPheIleGluSerLys	106	
DB	182	NN	241	
QY	107	ArgProArgLeuGluLeuLeuProAspProLeuLeuA-rgProSerProLeuLeuAlaThr	126	
DB	242	NN	301	
QY	127	GlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThrGlyLysLeuGlu	146	
DB	302	NN	361	
QY	147	ProValSerProProSerProProHisThrAspProGluLeuGluLeuValProProArg	166	
DB	362	CCGGTGTCTCCCGGACGCGCCGACACCGACCTGAGCTGAGCTGGTGGCGCGGG	421	
QY	167	LeuSerLysGluLeuLeuGlnAsnMetAspArgValAspArgGluIleThrMetVal	186	
DB	422	CTGTCCAAGAGGAGCTGATCAGAAACATGGACCGCGTGGACGAGAGATCAACATGTA	481	
QY	187	GluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGluGluAlaLys	206	
DB	482	GAGCAGCAGATCTTAAGCTGAAGACAGACGNNNNNNNNNTGAGGAGGAGCTWCCAAG	541	
QY	207	ProProGluProGluLysProValSerProProProIleGluSerLysHisArgSerLeu	226	
DB	542	CCGCGGAGCTGAGAAGCCGCTGTCAACCGCGCCCATCGAGTCAAGACCGCAGCCTG	601	
QY	227	ValGlnIleleThrAspGluAanArgLysLysAlaGluAlaAlaHisArgIleLeuGlu	246	
DB	602	GTGCAGATCATCTACGACGAGAACCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNN	561	
QY	247	GlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSerAspThrArgGlnTyr	266	
DB	662	NN	721	
QY	267	HisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIleLeuTyrPheLysArg	286	
DB	722	CATGAGAACATCAAAATAAACCGGCGATCGGAAGAAGCTAATCTTGACTTCAAGAGG	781	
QY	287	ArgAsnHisAlaArgLysGlnTyrLysGlnLysPheCysGlnArgTyrAspGlnLeuMet	306	
DB	782	AGGAATCACGCTCGGAACAAATGGAGCAGAGAGTCTGCCAGCGCTATGACAGCTCATG	841	
QY	307	GluAlaLeuGluLysLysValGluArgIleleuAsnAsnProArgArgAlaLysGlu	326	
DB	842	GAGGCTGGGAGAGAGGTTGAGCGCATCCAGAACCAACCCCGCGCGGCGCAAGGAG	901	
QY	327	SerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLysGlnArgGluLeu	346	

DB	902	AGCAAAGTGCAGAGTACTACGAGAAGCAGTTCCTGAGATCCGCAAGACGCGAGCTG	961	
QY	347	GlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeuSerMetSerAlaAla	366	
DB	962	CAGGAGCGCATGCAGAGCAGGTTGGCCAGCGGGGCACTGGTGTCTCCATGTCGGCTGC	1021	
QY	367	ArgSerGluHisGluValSerGluIleleAspGlyLeuSerGluGlnGluAsnLeuGlu	386	
DB	1022	CGCAGCGAGCAGAGGTGTCTCAGAGATCATCGATGCTCTCAGAGCAGAGAACCTGGAG	1081	
QY	387	LysGlnMetArgGlnIleAlaValIleProProMetLeuTyrAspAlaAspGlnGlnArg	406	
DB	1082	AAGCAGATCGCGAGCTGGCGTGTATCCCGCTATGCTGTACGACCGCCAGCAGCGC	1141	
QY	407	IleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLysValTyrLysAspArg	426	
DB	1142	ATCAAGTTTCATCAACATGAACGGCTTATGGCTGACCCCATGAAGGTGTACAAAGCCG	1201	
QY	427	GlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArgGluLysPheMetGln	446	
DB	1202	CAGGTCAATGAACATGTGAGTGAGCAGAGAGGAGACCTCCGGGAGAAANNNNNNNN	1261	
QY	447	HisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLysThrValAlaGluCys	466	
DB	1262	NN	1321	
QY	467	ValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSerLeuValArgSer	486	
DB	1322	GTCTCTATTACTTACTGACTAAGAAGAAATGAACTATAAGAGCCTGTGTAGACGAGC	1381	
QY	487	TyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	506	
DB	1382	TATCGCGCGCGGAGAGCCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1441	
QY	507	GlnGlnGlnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGluLysGlu	526	
DB	1442	NN	1501	
QY	527	LysGluAlaGluLysGluGluLysProGluValGluAsnAspLysGluAspLeuLeu	546	
DB	1502	NN	1561	
QY	547	LysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGluAlaAlaLaser	566	
DB	1562	NGGAGAGAGCAGACACCTCAGGGAGAGACCAACGACGAGAGGAGCGCGTCCCTCC	1621	
QY	567	LysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGlyArgIleThrArgSerMet	586	
DB	1622	AAAGCGCGCAAACTGCAACAGCAGGAGAGAGCGAAAGCGCCGATCACCCGCTCAATG	1681	
QY	597	AlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGlnSerAlaGluLeuAlaser	606	
DB	1682	GCTAATGAGGCAACAGCGAGGAGCCATCACCCCGCAGCAGCGCGAGCTGCCCTCC	1741	
QY	607	MetGluLeuAsnGluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGly	626	
DB	1742	ATGAGCTGATGAGAGTCTCGCTGGACAGAGAAATGGAACAGCAGCAAGAAAGGT	1801	
QY	627	LeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrVal	646	
DB	1802	CTCTCGAAACACGCGCGCACTGTCGCGCATCCCGGATGGTGGGCTCCAGACTGTG	1861	
QY	647	SerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeu	666	
DB	1862	TCGCGAGTGTAGAACTTCTTACTTCACTACAGAGAGGACAGACCTCGATGAGTCTG	1921	
QY	667	GlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysAla	686	
DB	1922	CAGCAGCAACAGCTCAAGATGAGAGAGAGAGAACCGCACGAGAGAGAGAGAGAGCG	1981	
QY	687	ProAlaAlaAlaSerGluGluAlaAlaPheProValValGluAspGluGluMetGlu	706	
DB	1982	CCGCGCGCGCGCAGCAGAGGCTGCTATTCGCGCCGCTGTGGAGGATGAGGAGATGAG	2041	

QY 707 AlaSerGlyValSerGlyAsnGluGluMetValGluGluAlaGluAlaLeuHisAla 726
DB 2042 CGCTCGGCGCTGAGCGGAAATGAGGAGGAGATGGTGGAGGAGCTGAAGNNNNNNNNNN 2101
QY 727 SerGlyAsnGluValProArgGlyGluCysSerGlyProAlaThrValAsnAsnSerSer 746
DB 2102 NNN 2161
QY 747 AspThrGluSerIleProSerProHisThrGluAlaAlaLysAspThrGlyGlnAsnGly 766
DB 2162 NNN 2221
QY 767 ProLysProProAlaThrLeuGlyAlaAspGlyProProGlyProProThrProPro 786
DB 2222 NNN 2281
QY 787 ArgArgThrSerArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaPro 806
DB 2282 NNN 2341
QY 807 ThrProProAlaProProSerProSerAlaProProProValProLysGluGlu 826
DB 2342 NNN 2401
QY 827 LysGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGlnLysProPro 846
DB 2402 NAGGAGGAGGAGACCGCAGCAGCGCCCGCAGTGGAGGAGGGGAGGAGCAGAAAGCCCC 2461
QY 847 AlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCys 866
DB 2462 GCGNNNGAGGAGCTGGCAGTGACACAGGAGGCGGAGGCGCGTCAAGAGCGAGTGC 2521
QY 867 ThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThr 886
DB 2522 ACGGAGGAGCGCGAGGAGGCGCGCGCAAGGCAAGGAGCGCGCGCTGAGGCTACG 2581
QY 887 AlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAla 906
DB 2582 GCCGAGGCGCGCTCAAGCGAGAGAAGAAGGAGGCGCGGAGCGGCGCAGCCACACAGCC 2641
QY 907 LysSerSerGlyAlaProGlnAspSerAspSerAlaThrCysSerAlaAspGluVal 926
DB 2642 AAGAGCTCGGCGCGCGCGCAGCAGCAGCAGCTCCAGTGCACCTGCGAGTGCAGACGAGTG 2701
QY 927 AspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuThr 946
DB 2702 GATGAGGCGGAGGCGCGGACAAAGAACCGGCTGTCTCCCAAGCGCGCAGCTCTCTACC 2761
QY 947 ProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLysGlnLeu 966
DB 2762 CCGACTGGCGACCCCGCGGCGCAATGCTCACCACAGAGCCACTGCACTGAAGCAGCTG 2821
QY 967 LysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProProArg 986
DB 2822 AAGCAGCAGCGGCTGCCATCCCCCATCCAGGTCCAGGTCCAGAAAGTCCATGAGCCCCCGG 2881
QY 987 GluAspAlaAlaProThrLysProAlaProProAlaProProProGlnAsnLeuGln 1006
DB 2882 GAGGACGAGCTCCCAAGCGCAGCTCCCGCAGCGCCCGCAGCCCAAGAACCTGCAG 2941
QY 1007 ProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerPro 1026
DB 2942 CCGGAGAGCGGCGCGCTCAGCAGCGCTGGCAGCAGCCCCCGGGGCAAGCAGGAGCGCG 3001
QY 1027 AlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspPro 1046
DB 3002 GCACCCCGCGCAGCAAGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3061
QY 1047 ProCysTTPThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSer 1066
DB 3062 NNN 3121

QY 1067 ProHisAlaProAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeu 1086
DB 3122 NNN 3181
QY 1087 GlyLeuHisAspThrAlaArgProValLeuProArgProProThrIleSerAsnProPro 1106
DB 3182 GGCCTCCATGACACTGCCCGCGCGCTCTGCGCGGCCACCCACCATCTCCAAACCGGCT 3241
QY 1107 ProLeuIleSerSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSer 1126
DB 3242 CCCCTTATCTCTCTGCCAGACACCCAGCGCTCTCGAGAGCAATAGGTGCAATCTCC 3301
QY 1127 GlnGlyMetSerValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGly 1146
DB 3302 CAAGAAATGTCGCTCCAGCTCCAGTCCGCTACTCAGAGCATGCCAAGCGCCCGTGGGC 3361
QY 1147 ProValThrMetGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGly 1166
DB 3362 CCGCTCACCATGGGCTGCCCTGCCATGACCCCAAAAGCTGGCACCTTNNCGCGA 3421
QY 1167 ValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyVal 1186
DB 3422 GTGAAGCAGGAGCAGCTGTCTCCACGGGGCCAGGCTGGGCCACCGAGAGCTGGAGTG 3481
QY 1187 ProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGly 1206
DB 3482 CCCACGGGCCAGGAGCATCTGCTGAGAGGACAGCTCTGGGCTCAGTTCCGGCGGA 3541
QY 1207 SerIleThrLysGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArg 1226
DB 3542 AGCATACCAAGGCAATTCACGAGCACACGGGTGCCCTCGGACAGCGCATCATACCGC 3601
QY 1227 GlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIle 1246
DB 3602 GGCTCCATCACCCACNNWACCCAGCTGCTCTGTANNNGGCCACCATCACAGGATC 3661
QY 1247 IleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGly 1266
DB 3662 ATCGCGAGGACAGCGNNNGTCGCTTGGANNCGCGCGGAGGACAGCTTGCCCAAGGC 3721
QY 1267 HisValIleTyrGluGlyLysGlyHisValLeuSerTyrGluGlyGlyMetSerVal 1286
DB 3722 CACGTCTACTCAGAGGCAAGGCGCACCTTGTCTATGAGGNNNNNNNNNNNNNNNN 3781
QY 1287 ThrGlnCysSerLysGluAspGlyArgSerSerGlyProProHisGluThrAlaAla 1306
DB 3782 NNN 3841
QY 1307 ProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSer 1326
DB 3842 NNN 3901
QY 1327 IleGluGlyLeuMetGlyArgAlaIleProProGluArgHisSerProHisIleLys 1346
DB 3902 NNNNNNGTCTCATGGCGCGTCCATCCCGCGGAGCGACAGCCCCCACCCTCAA 3961
QY 1347 GluGlnHisHisIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAla 1366
DB 3962 GAGCAGCACCATCCCGCGGTCCATCACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4021
QY 1367 GlnGluAspTyrLeuArgGluAlaLysLeuLysArgGluGlyThrProProPro 1386
DB 4022 NNN 4081
QY 1387 ProProProSerArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLys 1406
DB 4082 NNN 4141
QY 1407 LeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHis 1426
DB 4142 NNN 4201
QY 1427 GluIleProArgGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeu 1446

[illegible][illegible]

2167 ProGlyAlaSerCysProValLeuAspLeuArgProProSerAspLeuTyrLeuPro 2186
6422 NNN 6481
2187 ProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlySer 2206
6482 NNN 6541
2207 ProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSer 2226
6542 CCAGAGCCAAACAAGCGTCAGTCTTGGTGGCGGTGAGGACGGTATTGAACCTGTGTCC 6601
2227 ProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyr 2246
6602 CCACCGGAGGCGATCAGCGAGCAGCGGACCTCCCGAGTGTGTGTACCCGCTGTGTAC 6661
2247 ArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSer 2266
6662 CGGGATGGGGAACACAGCGAGCCACGAGGATGGGCTCCAAGTCTCCAGGCAACACGAG 6721
2267 GlnProProAlaPheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLys 2286
6722 CAGCCGCCAGCGCTCTTTCAGCAAGCTGACCGAGCAAACTCCGCCATGTCAAGTCCAG 6781
2287 LysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyr 2304
6782 AAGCAAGAGATCAACAAGAAGCTGAACACCCACACCGGAATGAGCGCTGATAC 6835

RESULT 4
BC033087
LOCUS
DEFINITION Homo sapiens, 2000 bp mRNA linear HTC 04-MAR-2003
IMAGE:5016291, mRNA.
ACCESSION BC033087
VERSION BC033087.1 GI:21619972
KEYWORDS HTC.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2000)
Strausberg,R.
Direct Submission
Submitted (25-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Boedet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Nees, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 43 Row: i Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5454073
This clone has the following problem: retained intron.
Location/Qualifiers

REMARK
COMMENT

source

1..2000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5016291"
/tissue_type="lung, small cell carcinoma"
/clone_lib="NIH MGC 7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Alignment Scores:
Pred. No.: 1.29e-96 Length: 2000
Score: 2566.00 Matches: 503
Percent Similarity: 90.72% Conservative: 15
Best Local Similarity: 88.09% Mismatches: 31
Query Match: 19.42% Indels: 22
DB: 11 Gaps: 1

US-09-522-753-5 (1-2517) x BC033087 (1-2000)

Qy 1 MetSerGlySerThrGlnLeuValAlaGlnThrTTPArgAlaThrGluProArgTyrPro 20
Db 273 ATGTCGGGATCCACACAGCCTGTGGCAGAGTGGAGGGCCACTGAGCCCGCTACCG 332
Qy 21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
Db 333 CCCACAGACCTTCTTACCCAGTGCAGATGCCCGGACGCACACGAGCGTGGGGTCTCTG 392
Qy 41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
Db 393 GAGTACCAGCACCACTCCCGGAGCTATGCTCCACCTGTCGCCGGCTCCATCATCCAG 452
Qy 61 ProGlnArgArgArgProSerLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
Db 453 CCCAGCGCGGAGGCGCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCAG 512
Qy 81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
Db 513 GAGTCCACCTGCGGCCAGAGTCCCATCTACCTGCCGAGCTGGGGAAGTCAGAGATG 572
Qy 101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
Db 573 GAGTTATTGAAGCAAGCGCCCTCGGTAGAGTGTGCTGACCCCTGTGTCGAGCG 632
Qy 121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
Db 633 TCACCCCTGTGCGCCACGCGGCTGCGGGATCTGAAGACCTTCAAGGACCGGTAGC 692
Qy 141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
Db 693 CTGAGCGCAAGCTGGAAACCGGTGTCTCCCCCAGCCCCCGCACACTGACCCCTGAGCTG 752
Qy 161 GluLeuValProProArgLeuSerLysGluLeuIleGlnAsnMetAspArgValAsp 180
Db 753 GAGCTGTGCGCCACAGCGCTGTCCAAGGAGAGCTGATCCAGAAATGACCGCGTGGAC 812
Qy 181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysGlnGlnLeu 200
Db 813 CGAGAGATCACCATGGTGTGAGCAGCAGATCTCTTAAGCTGGAAGAAAGACGCAACAGCTG 872
Qy 201 GluGluGluAlaAlaLysProGluProGluLysProValSerProProIleGlu 220
Db 873 GAGGAGGAGGCTGCGCAGCGCCCGGAGCTGAGAAAGCCCGTGTACCCGCCGCCATCAG 932
Qy 221 SerLysHisArgSerLeuValGlnIleTyrAspGluAsnArgLysLysAlaGluAla 240
Db 933 TCGAAGCACCAGCGCTGGTGCAGATCATCTACGACGAGAACCGGAGAGAGGCTGAAGCT 992
Qy 241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
Db 993 GCATCGGATTCTGGAGAGGCTGGGGCCCGAGGTGGAGCTGCCGCTGTACAACACGAGCC 1052

QY	261	SerAspThrArgGlnTyrHisGluAsnIleLysLeuAlaMetArgLysLysLeu	280
DB	1053	TCCGCACCCCGCAGTATCATGAGAACAATCAAAATAAACACGAGCGATTCGGGAAGAAGCTA	1112
QY	281	IleLeuTyrPheLysArgArgAsnHisalaargLysGlnTrpLysGlnLysPheCysGln	300
DB	1113	ATCTTGTAATTCAAAGAGGAGGAATCACGCTCGGAACAACATGGGAGCAGAGATTCTGCCAG	1172
QY	301	ArgTyrAspGlnLeuMetGluAlaLeuGluLysValGluAArgIleGluAsnAsnPro	320
DB	1173	CGTTATGACCAGGTCAATGAGGCCTGGGAGAGAGAGTGGAGCGCATCGAACAACCCC	1232
QY	321	ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle	340
DB	1233	CGGCGCGGCCCAAGGAGAGCAAGGTGCGCAGTAGTACTACGAAGACAGTTCCTCTGAATC	1292
QY	341	ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly	360
DB	1293	CGCAAGCAGCGCGAGCTGCGAGGAGCGCATGCAAGCAGAGGTGGCCAGCGGGCAGTGGG	1352
QY	361	LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer	380
DB	1353	CTGTTCATGTGCGCGCGCCGACGAGCGACGAGGTGTCAAGATCATCGATGGCCTCTCA	1412
QY	381	GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProMetLeuTyr	400
DB	1413	GAGCAGGAGAACCTGGAGAGACGATGCGCCAGCTGGCCGTGATCCCGCCCATCTGTATC	1472
QY	401	AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet	420
DB	1473	GACGCTGACACGACGAGCGCATCAAGTTCATCAAATGACGGGCTTATGGCCGACCCCATG	1532
QY	421	LysValTyrLysAspArgGlnValMetAsnMetTyrSerGluGlnGluLysGluThrPhe	440
DB	1533	RAGGTGTCAAAAGACCGCCAGTTCATGAACATGTGGAGTGAGCAGGAGAGGAGGACCTTC	1592
QY	441	ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg	460
DB	1593	CGGGAGAAGTTTCATGCACGATCCCAAGAAGCTTTGGCCTGATCGCATCATGAGAACTATAAG	1712
QY	461	LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys	480
DB	1653	ALAGACAGTGGCTGAGTGGCTCTTACTTACTACCTGACTAAGAAGAATGAGAACTATAAG	1712
QY	481	SerLeuValArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGln	500
DB	1713	AGCCTGGTGAGACGGAGTATCGCGCGCGCGCAGAGCCTGCTGCTGTATTAAGTTCCAG	1772
QY	501	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGluLys	520
DB	1773	CAAAAAGAATCACCGCGCGAGTAAAGAGTCCAGGCCAGAGAGTTCCTGCCCCAGAAAG	1832
QY	521	AspGluLysGluLysGluAlaGluGluGluLysGluGluLysProGluValGluAsn	540
DB	1833	CCACAGGCCAGAAAGCAGCGCTGTCTCAAAAAGCTCAGAGGGTCAAAAAGCTCCAGCCC	1892
QY	541	AspLysGlu-----	543
DB	1893	AGAAAAGCACCTGTCTCCAAAAGGCATTCGGCAAGAAAGCATAGTGGCAATCATAAAAAGTA	1952
QY	544	-----AspLeuLeuLysGluLys 549	
DB	1953	ATAAAGGTTCTTTTGACCTGTATAAAAAAAAAA 1985	
<hr/>			
RESULT 5			
BC020427			
LOCUS			
DEFINITION Homo sapiens, clone IMAGE:4179307, mRNA.			
ACCESSION BC020427			
VERSION BC020427.1 GI:18043012			
KEYWORDS HTC.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
linear HTC 03-JAN-2002			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1362)

Strausberg, R.

Direct Submission

Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalon@bcm.tmc.edu

Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gunaratne, P., Yoon, V., Kowis, C., Martin, R., Lawrence, S., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>

Series: IRAK Plate: 12 Row: i Column: 23

This clone has the following problem: no polyA-tail.

Location/Qualifiers

1. .1362

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4179307"

/tissue_type="Brain, anaplastic oligodendroglioma with 1p/19q loss"

/clone_lib="NCI CGAP_Brn67"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.: 2,7e-84 Length: 1362

Score: 2264.00 Matches: 431

Percent Similarity: 99.54% Conservative: 0

Best Local Similarity: 99.54% Mismatches: 2

Query Match: 17.13% Indels: 0

DB: 11 Gaps: 0

US-09-522-753-5 (1-2517) x BC020427 (1-1362)

QY 1676 ThrAlaAlaLeuGluAsnArgGinThrIleleAsnAspTyleThrSerGlnGlnMet 1695

Db 63 ACGCGGGCTGGAGACCGGCAGACCATCATCAATCACTATCATCACTCGAGAGATG 122

QY 1696 HisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerPro 1715

Db 123 CACCACAAACGGGGCCACCGGCATGGCCAGCGAGCTGATATGCTGAGGGGCTCTCGCC 182

QY 1716 ArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleleAspLeuSer 1735

Db 183 CCGGAGTCTCTGCTGGCACTCAACTACGCTGGGGTCCCCAGGCATCATCGACCTGTCC 242

QY 1736 GlnValProHisLeuProValLeuValProThrProThrProGlyThrProAlaThrAlaMet 1755

Db 243 CAAAGTGCCACACTCTGCTGTCTCGTGCCTCCGACACACAGGCACCCAGCCACCGCATG 302

QY 1756 AspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerSer 1775

Db 303 GACCGCTTGCTACTACTCCCAACCGGCCCCAGCCCTTCAGCAGCGCCACAGCAGCTCC 362

QY 1776 ProLeuSerProGlyGlyProThrHisLeuThrIysProThrThrThrSerSerGlu 1795

572	CCAGGTTCTGCAGGCGCTCAAGAA-----AGGAGAACTAGTTATGAAACGGTTTCATCA	625
94	GlueLgUlySerGluMetCluPheIleGluSerLysArgProArgLgUleuGluLeuLeu	113
626	GGCCCATCCCGAGTGATCATGATCACTGGAATCGAAGCGACCAAGCTCTGGAAACAGGTT	685
114	ProAsp-----ProLeuLeuArgProSerPro	122
686	TCGTGATCTCATTTTCAGCGTGTCAAGTGTGCGGTTTTCGCTTTAGTGCACCGCGTCCCA	745
123	LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr	142
746	-----GAAGGGCTGAGGCT--TCTGCAGATGCTAAGACAGATCCAGCATTCGGA	793
143	GlyLysLeuGlu--ProValSerProProSerProProHisThrAspProGluLeuGlu	161
794	GGCAACATGAAAGCTCCATCTCTCCAAATTCGGGGCAACCATGTGGAGATGATCAAAAT	853
162	LeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArg	181
854	GCTTCACCTTCAAAACCTCTCAAAGGAAGAGTTAAATACAGAGTATGGATCGTGTAGATCGA	913
182	GluIleThrMetValGluGlnGlnIleSerLysLeuLysLysGlnGlnGlnLeuGlu	201
914	GAATTCGAAAGATGAGACAGCATCTTAAACTGAAAAAGAAACAAACACAGCTTGAA	973
202	GluGluAlaLysProProGluProGluLysProValSerProProProIleGluSer	221
974	GAAGAGGCGAGTAAACCTCTCGAGCTCGAAGCCCGTGTCCCTCTCTCTGTGGCGAG	1033
222	LysHisArgSerLeuValGlnIleLeuTyAspGluAsnArgLysLysAlaGluAla	241
1034	AAACACCCAGTATTTGCCAAATTAATTTATGATGAGATCGGAAAAAGCAGAGAAGCT	1093
242	HisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyAsnGlnProSer	261
1094	CATAAAATTTTGAAGGCTTTGGGCCCAAAAGTTGAACTGCCACTGTATTAACCGCATCA	1153
262	AspThrArgGlnTyHisGluAsnIleLysIle-----	272
1154	GATACCAAGGTGACCATGAGAACATCAAGACTGGAGTACCTCAAGGGCGCATGATAA	1213
273	AsnGlnAlaMetArgLysLysLeuIleLeuTyPheLysArgArgAsnHisAlaArgLys	292
1214	AAACAGGTGATCAGGAAAAAACTCATTTTATTTTAAAAAGAGAATAATCATGCAAGAAA	1273
293	GlnTyLysGlnLysPheCysGlnArgTyAspGlnLeuMetGluAlaLeuGluLysLys	312
1274	CAAAGGGAACAAAAATCTGCCAGCTTATGATCAGCTCATCGAGGCGATGGGAGAAAAA	1333
313	ValGluArgIleGluAsnAsnProArgArgAlaLysGluSerLysValArgGluTy	332
1334	GTGCACAGATAGAAAAATAATCCTCGAGGAAAGCTAAAGAAAGCAAAACAGGGGAATAC	1393
333	TyrGluLysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSer	352
1394	TATGAAAGCGATTCAGAAATTCGAAACCAAGAGAACACAGCAAGAAAGATTTCAG--	1450
353	ArgValGlnGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluVal	372
1451	CGAGTTGGGCGAGGGGAGCTGGTCTTTTCAGCCACCATTTGTAGAGTGTAGCATGAGATT	1510
373	SerGluIleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeu	392
1511	TCTGAAATATTGATGGGTCTCTGAGCAGGAGAAATATATGAGAACAAATATCGGCAGCTC	1570
393	AlaValIleProProMetLeuTyAspAlaAspGlnArgIleLysPheIleAsnMet	412
1571	TCTGTGATTCACCTGATGTTTGTATGAGAAACAAACAGCAGTCAAGTTCATTAACATG	1630
413	AsnGlyLeuMetAlaAspProMetLysValTyLysAspArgGlnValMetAsnMetTyr	432
1631	AATGGGCTTATGGAGGACCTATGAAAGTGTATTAAGATATGACAGTTTATGATGTTGG	1690

Qy	433	SerGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysIlePheGly	452
Dd	1691	ACTGACCATGAAGAAGAGATCTTTAAGGACAAGTTCATCCAGCATCCAAAAAACCTTGGA	1750
Qy	453	LeuILEalaserPheLeuGluArgLysThrValalaGluCysValLeuTVTYrTyLeu	472
Dd	1751	CTAATTGCATCATCACTTGGAGAGAAAGGTTCCTGATTGTGTTTTGTTATTAATTATTA	1810
Qy	473	ThrLylysAsnGluAsnTyrLysSerLeuValArgSerTyr---	491
Dd	1811	ACCAGAAGAAAATGGAATTATAAGCCCTCGTCAGAAGGAATATATGGAAACGACGAGGC	1870
Qy	492	LysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	511
Dd	1871	AGAAACACAGCAA-----	1882
Qy	512	MetProArgSerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGluLys	531
Dd	1883	ATTGCTCGACCTCGCAGAGAGAAAAGTAGAAGAAAAGAGAGGATAAAGCAGAGAAAAA	1942
Qy	532	GluGluGluLys	535
Dd	1943	ACAGAAAAAANA 1954	
RESULT 7			
BC035748		1950 bp mRNA linear	HTC 19-NOV-2003
LOCUS		Homo sapiens nuclear receptor co-repressor 1, mRNA (cdna clone	
DEFINITION		IMAGE:5749816), with apparent retained intron.	
ACCESSION		BC035748	
VERSION		BC035748.1 GI:23958885	
KEYWORDS		HTC.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
		1 (bases 1 to 1950)	
		Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,	
		Klautner,S.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,	
		Altshuler,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,	
		Hopkins,R.F., Jordan,H., Moore,T.T., Max,A.M., Wang,J., Hsieh,F.,	
		Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,	
		Casavant,T.L., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,	
		Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,	
		Carninci,P., Prange,C., Roh,S.S., Loquellano,N.A., Peters,G.J.,	
		Abramson,R.D., Mullaly,S.J., Bosak,S.A., McSwan,P.J.,	
		McKernan,K.J., Mallek,J.A., Gunaratne,P.H., Richards,S.,	
		Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,	
		Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,	
		Fahney,J., Helton,E., Ketterman,M., Madan,A., Young,A.C., Shevchenko,Y.,	
		Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,	
		Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,	
		Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalus,D.E.,	
		Schmerer,A., Schein,J.E., Jones,S.J. and Marra,M.A.	
		Generation and initial analysis of more than 15,000 full-length	
		human and mouse cDNA sequences	
		Nature Genet. 1999 Sep; 21(2): 99 (26). 16899-16903 (2002)	

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL	REMARK COMMENT
--	-------------------

USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-i@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.

QY	92	uProGluLeuGlyLysSerGluMetGluPhe-----IleGluSerLysArgProArgLeu	110
Db	517	-----GGACCCCTCAGCGGTGGATCATGATTCCTTGGAGTCACAGGCCCGCCAGCCCT	566
QY	110	uGluLeuLeuProAsp-----ProLeuLeuAr	119
Db	567	GGAGCCAGTTCTCATGCCCATTTCCAGCGGTGTAGTGTGGCTTTTACCTTTAGTTCA	626
QY	119	gProSerProLeuLeuAlaThrGlyGlnProAlaGly-----SerGluAspLeuTh	136
Db	627	C-----TCGCTCCAGAGAGGCTTCAGTGTGCTGCATGCTAA	665
QY	136	ryAspArgSerLeuThrGlyLysLeuGlu---ProValSerProProSerProProHi	155
Db	666	GAGAGATTTCAGCATTTGGAAGCAACATGAGCTCCATCTCTCTTGTGGCTGGCAACC	725
QY	155	sThrAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuLeuGlnAs	175
Db	726	ATGTGGAGATGACCAAAATGCTTCACCTTCACAGCTTCAAGAGGAGGATTAATACAGAG	785
QY	175	nMetAspArgValAspArgGluLeuThrMetValGluGlnGlnIleSerLysLeuLysLy	195
Db	786	TATGACCGGTAGACCGGAGAGATTGCAAAAGTAGAGCAGCATCTCTTAATTTGAAAA	845
QY	195	elysGlnGlnGlnLeuGluGluAlaAlaLysProProGluProGluLysProValSe	215
Db	846	GAACAAACAAAGCTAGAGAGAGAGAGCAGCTAAAGCCCTCAACAGAGAGCCCTGTGTC	905
QY	215	rProProPheGluSerLysLysArgSerLeuValGlnIleLeuThrAspGluAsnAr	235
Db	906	CCCTCTCTCTGTGGAAACAAACACCGTAGTATTGTCCTCAAAATTTATGATGAGATCG	965
QY	235	glyLysAlaGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	255
Db	966	GAAG	1025
QY	255	oLeuTyAsnGlnProSerAspThrArgGlnTyHisGluAsnIleLysIleAsnGlnAl	275
Db	1026	ACTCTATAACCCATCATGATACCAAGAGTGTACCATGAGACATCAAGACAAACAGGT	1085
QY	275	aMetArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys	295
Db	1086	GATGAG	1145
QY	295	sGlnLysPheCysGlnArgTyAspGlnLeuMetGluAlaLeuGluLysLysValGluAr	315
Db	1146	ACAAAAATCTGCCAGCGCTATGATCAGCTCATGGAAGCGTGGGAGAAAAAAGTGACAG	1205
QY	315	gileGluAsnAsnProArgArgAlaLysGluSerLysValArgGluTyTyGluLy	335
Db	1206	AATAGAAAAATATCCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1265
QY	335	sGlnPheProGluLeuArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGl	355
Db	1266	GCAGTTTCCAGAAATTCGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1322
QY	355	yGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIl	375
Db	1323	TCAGAGGGGAGTGTCTTTTCAGCCACCATTCGTAGAGTGTAGAGATTTCTGAAT	1382
QY	375	eileAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIl	395
Db	1383	TATTGATGTCTTCTGAAACAGGAGATTAATGAGAGCAATGCGTCAGCTTCTGTGAT	1442
QY	395	eProProMetLeuTyAspAlaAspGlnGlnArgLysPheIleAsnMetAsnGlyLe	415
Db	1443	TCCACCTATGATGTTTGTAG	1502
QY	415	uMetAlaAspProMetLysValTyLysAspArgGlnValMetAsnMetTrpSerGluGl	435
Db	1503	GATGAGGAGATCAATGAAGAGTTTATAAAGACACAGCTTTATGAATGTTGGACTGACCA	1562

QY	435	nGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuLeAl	455
Db	1563	TGAAGAGAGATCTTTAAGGACAGATTTATCCAGCATCCAAAAAATTTGGACTAATTCG	1622
QY	455	asePheLeuGluArgLysThrValAlaGluCysValLeuTyTyTyLeuThrLysLy	475
Db	1623	ATCCTATTGGAAAGGAAGAGTGTCTGATGTGTGTGTATATATATATATATATATATAT	1682
QY	475	sAsnGluAsnTyLysSerLeuValArgArgSerTyT---ArgArgArgGlyLysSerGl	494
Db	1683	AAATGAGAATTATAGGCCCTCGTGAGAGGAATTAATGGAAGAGCAGAGGAGAGATCA	1742
QY	494	nGln	514
Db	1743	GCAGCAG-----ATTGCCCG	1757
QY	514	gSerSerGlnGluLysAspGluLysGluLysGluLysGluLysGluLysGluLysGluGl	534
Db	1758	TCCCTCACAAGAGAAAAAGTAGAGAAAGAGAGATTAAGCAGAGAAAAAANA	1817
QY	534	u 534	
Db	1818	A 1818	
RESULT 10			
BM909096			
LOCUS			
DEFINITION			
5', mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. 1093			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:5431719"			
/tissue_type="astrocytoma grade IV, cell line"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH MGC 98"			
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
6.68e-58			
1624.00			
Length:			
Matches:			
1093			
331			


```

Db      550  GAAATTCGAGTCTCTGACTACCGCTCTCTCATATTGAAGTTAGCCAGCATCACAGCTT 609
Qy      54   SerProGlySerLeuIleGlnProGln---ArgArgArgProSerLeuLeuSerGluPhe 72
Db      610  TTACAGCAGCAGCAGCAGCAGCAGCTTCGAAGACGACCTTCCTTGCTTTCAGAAATTT 669
Qy      73   GlnProGlyAenGluArgSerGlnGluLeuHisLeuArgProGluSer---HisSerTyr 91
Db      670  CACCGGGTCTGACAGCCTCAAGAGGAGGAGAACTGGATATCAACAGATTTCACTCA--- 726
Qy      92   LeuProGluLeuGlyLysSerGluMetGluPhe-----lLeGluSerLysArgProArg 109
Db      727  -----GGACCTCCTACCGGTGATCATCATCTTCCTGGAGTCAAGCCGCCACGC 774
Qy      110  LeuGluLeuLeuProAsp-----ProLeuLeu 118
Db      775  CTGGAGCCAGTTCTGTATGCCATTTCCAGCGTGTAGTGTGGGGTTTACCTTTAGTT 834
Qy      119  ArgProSerProLeuLeuAlaThrGlyGlnProAlaGly-----SerGluAspLeu 135
Db      835  CAC-----TCGCTGCCAGAGGCTTGAGTGTCTGCGATGCT 873
Qy      136  ThrLysAspArgSerLeuThrGlyLysLeuGlu---ProValSerProProSerPro 154
Db      874  AAGAAGGATTCAGCATTTTGGAGCAAAACATGAAGCTCCATCTCTCTTGGCTGGCGAA 933
Qy      155  HisThrAspProGluLeuGluValProArgLeuSerLysGluGluLeuIleGln 174
Db      934  CCATGTGGAGATGACCAAAATGCTTCACTTCAAAAGCTTTCAAAGGAGGATTAATACAG 993
Qy      175  AsnMetAspArgValAspArgGluLeuThrMetValGluGlnGlnIleSerLysLeuLys 194
Db      994  AGTATGACCGGTAGCAGAGATTCGAAAAGTAGAGCAGCATCTTAAATTTGAAA 1053
Qy      195  LysLysGlnGlnLeuGluGluAlaLysProGluProGluProGluLysProVal 214
Db      1054  AAGAAACAACACAGCTAGAGAGAGCAGTGAAGCCCTTGAACACGAGAGGCTGTG 1113
Qy      215  SerProProGluSerLysHisArgSerLeuValGlnIleLeuThrAspGluLeu 234
Db      1114  TCCCTCTCTCTGTGAACAAACACACCCGTAGTATTGTCCAAATTTATTTATGAGAAAT 1173
Qy      235  ArgLysLysAlaGluAlaAlaHisArgLysLeuGluGlyProGlnValGluLeu 254
Db      1174  CGGAAAACAGCAAGAGCTCAATAATATTGAAGTCTTGGCCCAAAAGTTGAACGT 1233
Qy      255  ProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnLleLysLysLeuGln 274
Db      1234  CCATCTATAACCGCCATCAGATACCAAGGTGTACATGAGAACATCAAGACAAACCCAG 1293
Qy      275  AlaMetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrp 294
Db      1294  GTGATGAGGAAAAAACAACCTCAATTTATTTTAAAGAGAAATCATCAAGAAACAAAGG 1353
Qy      295  LysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGlu 314
Db      1354  GAAACAAAATCTCCACCGCTATGATCAGCTCATGGAAGCGTGGGAGAAAAGTGGAC 1413
Qy      315  ArgLleGluAsnAsnProArgArgAlaLysGluSerLysValArgGluTyrTyrGlu 334
Db      1414  AGAATAGAAAATAATCCCGAGGAAAGCAAAAGAAAGCAAAACAGGGAAATATCTATGAG 1473
Qy      335  LysGlnPheProGluIleLeuLysGlnArgGluLeuGlnGlnArgMetGlnSerArgVal 354
Db      1474  AAGCAGTTTCCAGAAATTCGAAAACAAAGAGACAGCAAGAAATTCAG---CGAGTT 1530
Qy      355  GlyGlnArgGlySerGlyLysSerMetSerAlaAlaArgSerGluHisGluValSerGlu 374
Db      1531  GGTACAGGGGAGCTGGTCTTTCAGCCACCATTCGTAGGAGTGCAGATGAGATTTCTGAA 1590
Qy      375  llelleAspGlyLeuSerGluGluAsnLeuGluLysGlnMetArgGlnLeuAlaVal 394

```

```

Db      1591  ATTATTGATGGTCTTTCTGAACAGGAGAAATATGAGAAAGCAATGGCTCAGCTTTCTGTG 1650
Qy      395  ileProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGly 414
Db      1651  ATTCACCTATGATGTTTGTATGTCAGAACAAAGAGGGTCAAAATTCATCAATATGAATGGG 1710
Qy      415  LeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGlu 434
Db      1711  CTGATGGAGGATCCAAATGAAGGTTTATAAGACACACAGAGTTTATGAATGTTTGGACTGAC 1770
Qy      435  GlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIle 454
Db      1771  CATGAAAAGGAGATCTTTAAGGACAGAGTTTATCCAGATCCAAAAAACTTTGGACTAAT 1830
Qy      455  AlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrLeuThrLys 474
Db      1831  GATCTCTATTGGAAGGAGAGTGTCTCTGATTGTGTGTTTATATTTACTATTAAACCAAG 1890
Qy      475  LysAsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArgGlyLysSer 493
Db      1891  AAAAATGAGATTAAGAAGCCCTCGTGAAGGAATATGGAACCGCAGAGCGAGAAAT 1950
Qy      494  GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 513
Db      1951  CAGCAGCAG-----ATTGCC 1965
Qy      514  ArgSerSerGlnGluLysAspGluLys 523
Db      1966  CGTCCCTCACAGAGAGAAAAAAGAAAAA 1995

RESULT 12
BM555371 965 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6577786 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470428
DEFINITION 5', mRNA sequence.
ACCESSION BM555371 GI:18795772
VERSION BM555371.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 965)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LICM1977 row: d column: 13
High quality sequence stop: 670.
FEATURES
Location/Qualifiers
1..965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5470428"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 41"
/note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a

```

NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 2,86e-55 Length: 965
Score: 1559.00 Matches: 304
Percent Similarity: 95.00% Conservativeness: 0
Best Local Similarity: 95.00% Mismatches: 11
Query Match: 11.80% Indels: 5
DB: 12 Gaps: 3

US-09-522-753-5 (1-2517) x BM553371 (1-965)

QY 2182 AspLeuTyrLeuProProAspHisGlyAlaProAlaArgGlySerProHisSerGlu 2201
Db 3 GACCTCTACCTCCCGCCCGGACCATGGTGCCTCCCGCGGTGGCTCCCGCCACACGCA 62
QY 2202 GlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGluAspGly 2221
Db 63 GGGGCAAGAGGTCTCCAGAGCAACAGACGTCGCTTTGGGTGGTGGAGACGGT 122
QY 2222 IleGluProValSerProGluGlyMetThrGluProGlyHisSerArgSerAlaVal 2241
Db 123 ATTGAACCTGTCTCCCGAGGCGCATGACGGAGCCAGGCATCTCCGGAGTCTGTG 182
QY 2242 TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSer 2261
Db 183 TACCGCTGTGTACCGGATGGGGAACAGACGAGCCAGCAGGATGGCTCCAAAGTCT 242
QY 2262 ProGlyAsnThrSerGlnProProAlaPheSerLysLeuThrGluSerAsnSerAla 2281
Db 243 CCAGGCAACACCCAGCCGCGCCCTCTTTCAGCAAGCTGACCGAGCAACTCCGCC 302
QY 2282 MetValLysSerLysGlnGluLeuLeuLysLysLeuAsnThrHisAsnArgAsnGlu 2301
Db 303 ATGGTCAAGTCCAGAGCAAGAGATCAACAGAAAGCTGAACACCACCAACCGGAATGAG 362
QY 2302 ProGluTyrAsnLysSerGlnProGlyThrGluLysPheAsnMetProAlaLysThrGly 2321
Db 363 CCTGAATCAATATCAGCCAGCTGGGAGGAGATCTTCAATATGCCCGCCATCACCGGA 422
QY 2322 ThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGly 2341
Db 423 ACAGGCTTATGACCTATAGAGCCAGCGGTGTCAGGAACATGCCAGCACCAACATGGGG 482
QY 2342 LeuGluAlaLysLysArgLysLeuMetGlyLysTyrAspGlnThrGluGluSerPro 2361
Db 483 CTGGAGGCCATAATTAGAAAGGCATCTATGGGTAAATATGACCATGGGAAGAGTCCCGG 542
QY 2362 ProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMet 2381
Db 543 CGGCTCAGGCAATGCTTTTAACCTCTGAATGCCAGTGCAGCGCTGCCCGCTGCTATG 602
QY 2382 ProLeuThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2401
Db 603 CCCATAACCGCTGTCTGACGAGCGAGTGAACACACACTCACCTGCCAGGTGGCGCGGG 662
QY 2402 LysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeu 2421
Db 663 AAGGCCAAGGTCTCTGGCAGACCCAGCAGCGCAAAAGCCAAAGTCCCGCGCCCGCGCTG 722
QY 2422 AlaSerGlyAspArgProSerValSerSerValHisSerGluGlyAspCysAsnArg 2441
Db 723 GCATCTGGGACCGCCACCTCTCTCTCTCAGTGCACTCGGAGGAGAGTGCACACCGC 782
QY 2442 ArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrPro 2461
Db 783 CGGACCGCGCTCACCAACCGGTGTGGGAGGACAGCGCTCTGTCGCGAGGTGCCACGCCA 842
QY 2462 PheProTyrAsnProLeuLysMet-ArgLeuGlnAlaGly---ValMetAlaSerProPr 2480
Db 843 TTCCCTCAAAACCCCTCATCATGNCNGTGCAGGCGGGNTGGTCATGGTTTTCGCCACC 902

QY 2480 oProProGly---LeuProAlaGly-----SerGlyProLeuAlaGlyProHisHis 2496
Db 903 CCCACCGGGGCTTCCCGCGGGGAGCGGGGGCCCCCTCGGCTGGGGCCCCCACC 960

RESULT 13

BM553310 1126 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6572642 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467163
5', mRNA sequence.

ACCESSION BM553310

VERSION BM553310.1 GI:18791930

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1126)

NIH-MGC <http://mhc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/BTP

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing: The I.M.A.G.E. Consortium (LLNL)

Distributing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LUCM1968 row: 1 column: 12

High quality sequence stop: 674.

FEATURES

source

1..1126

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="IMAGE:5467163"

/tissue_type="amelanotic melanoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_41"

/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCACAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 2,08e-54 Length: 1126
Score: 1540.00 Matches: 314
Percent Similarity: 82.19% Conservativeness: 9
Best Local Similarity: 79.90% Mismatches: 16
Query Match: 11.65% Indels: 55
DB: 12 Gaps: 1

US-09-522-753-5 (1-2517) x BM553310 (1-1126)

QY 1457 LeuLysTyrAspThrGlyAlaSerThrGlySerLysLysHisAspValArgSerLeu 1476

Db 2 CTCAGTACGACACCGCGCGCTCCACCCTCCCAAAAGACGACGCTACGCTCCCTC 61

QY 1477 IleGlySerProGlyArgThrPheProValHisProLeuAspValMetAlaAspAla 1496

Db 62 ATCCGACGCCCGCGCGAGCTTCCACCCGTCGACCCGCTGGATGTATGCGCGACGCC 121

QY 1497 ArgAlaLeuGluArgAlaCysTyrGluSerLeuLysSerArgProGlyThrAlaSer 1516

Db 122 CGGCACCTGGANCTGCTCTACGAGAGAGCTGAAGAGCCGCGCAGGACCCCGAC 181

QY 1517 SerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysPro 1536

Db	182	AGCTCGGGGCTCCATTGCGCGGGCCCGCTCATTTGCTGAGCTGGGCAAGCCG	241	VERSION	BM477568.1	GI:18526610
Qy	1537	ArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArg	1556	KEYWORDS	EST.	Homo sapiens (human)
Db	242	CGGCAGAGCCCTTAACCTATGAGACACACGGGACACCTTTGCCGGCACCCTCCACGA	301	SOURCE	Homo sapiens	
Qy	1557	GlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSer	1576	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	302	GGTTTCGGCCGTCACACCGCGGAGCCACCGCGCTGAGAGGGCAGCTTTCGTCC	361	REFERENCE	1. (bases 1 to 1057)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Qy	1577	SerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluLleAlaLysSerPro	1596	AUTHORS	NIH-MGC http://mgi.nci.nih.gov/	
Db	362	AGCAAGGATCCACGACCGAAGTACGTCGACGCTGTGAGATGCGCAAGTCCCG	421	TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
Qy	1597	HisSerThrValProGluHisProHisProLleSerProTyrGluHisLeuLeuArg	1616	JOURNAL	Unpublished (1999)	
Db	422	CACGACCGTCCCGAGACACCCACACCCATCTGCCCTATGACACCTGCTTCGG	481	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs@mail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Qy	1617	GlyValSerGlyValAspLeuTyrArgSerHisLleProLeuAlaPheAspProThrSer	1636	FEATURES	Plate: LLAM12272 row: p column: 05 High quality sequence stop: 674.	
Db	482	GGCGTGAGTGGGTGACCTGATCGCAGCCACATCCCTCGGCTTCGACCCACCTCC	541	source	Location/Qualifiers	
Qy	1637	IleProArgGlyLeuProLeuAspAlaAlaAlaTyrTyrLeuProArgHisLeuAla	1656	1. .1057	/organism="Homo sapiens"	
Db	542	ATACCCCGCGGATCCCTCTGACGACGCGCTGCTACTACTACCTGCGCCGACACTGGCC	601	/mol_type="mRNA"	/db_xref="taxon:9606"	
Qy	1657	ProAsnProThrTyrProHisLeuTyrProTyrLeuLleArgGlyTyrProAspThr	1676	/clone="IMAGE:5554420"	/tissue_type="lymphoma, cell line"	
Db	602	CCCAACCCACCTACCCGACCTGTACCCACCTACCTACCTACCTACCTACCTACCTAC	661	/lab_host="DH10B (phage-resistant)"	/notes="Organ: lymph; vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
Qy	1677	AlaAlaLeuGluAsnArgGlnThrLleLeuAsnAspTyrLleThrSerGlnGlnMethis	1696	ORIGIN		
Db	662	CGCGCGCTCGAGAACCGGACGACCATCATATGATCATCATCATCATCATCATCATCAT	721	Alignment Scores:		
Qy	1697	HisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArg	1716	Pred. No.:	5.63e-52	Length: 1057
Db	722	CACAACCGCGCCACCGCCATGCGCCGAGCGCTGATGCTGAGGGGCTCTGCGCCCGC	781	Score:	1480.50	Matches: 308
Qy	1717	GluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyLleLleAspLeuSerGln	1736	Percent Similarity:	87.57%	Conservative: 2
Db	782	GAGTCTCGTGGCACTCAACTAGCGTGGGTGCGGCTCCCGA	820	Best Local Similarity:	87.01%	Mismatches: 27
Qy	1737	ValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAsp	1756	Query Match:	11.20%	Indels: 20
Db	820	-----	820	DB:	12	Gaps: 6
Qy	1757	ArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerPro	1776	US-09-522-753-5 (1-2517) x BM477568 (1-1057)		
Db	820	-----	820	Qy	929	AlaGluGlyGlyAspLysAsnArgLeuSerProArgProSerLeuLeuThrProThr
Qy	1777	LeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArg	1796	Db	1	GCCAGGGCGGCGCAAGAACCGGCTGTCTCCCAAGGCCAGCTCTCTCACCCGACT
Db	821	-----	820	Qy	949	GlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLysGlnLeuLysGln
Qy	1797	GluArgAspArg-AspArgGluArgAsp-ArgAspArgGluArgGluLysSerLleLeuT	1816	Db	61	GCGACCCCGCGGCAATGCTCACCCAGAACCACTGGACCTGAAGCAGTGAAGCAG
Db	871	GAGCCAGACCGGGATCCAAAGCGGGAACCGGAATTCGACGCGGGAAGTCCATCTCC	930	Qy	969	ArgAlaAlaLleProProLleGlnValThrLysValHisGluProProArgGluAsp
Qy	1816	hrSerThrThrVal-GluHisAlaProLleTyr-ArgProGlyThrGluGlnSerSe	1835	Db	121	CGAGCGGTGCTCATCCCCCATCCAGTCAACAAAGTCCATGAGCCCCCCCCCGGAGAC
Db	931	CGTCCACACCGAGTGGAGCCGACCCCTCTGGAAACCTGTAAACGAGCAACAA	990	Qy	989	AlaAlaProThrLysProAlaProProAlaProProProGlnAsnLeuGlnProGlu
Qy	1835	rgLysSerSerGlySerSerGlyGlyGly 1845		Db	181	CGAGTCTCCACAGCCAGCTCTCCCGACCCCGCCAGCCGCAAAACCTGCGACCGGAG
Db	991	CGGAAGCATCCGGGAAACTGGGGGGGGGT 1021		Qy	1009	SerAspAlaProGlnGlnProGlySerSerProArgLysSerArgSerProAlaPro
RESULT 14				Db	241	AGCAGCGCCCTCAGCAGCTGGCAGCAGCCCCCGGGCAAGCAGGAGGCCCGGACCC
BM477568				Qy	1029	ProAlaAspLysGlu-----AlaPheAlaAlaGluAlaGln
LOCUS				Db	301	CCCGCCGCAAGGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DEFINITION				Qy	1041	LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg
5', mRNA sequence.				Db		
ACCESSION						

```

Db 361 AAGTCCTGGGAGACCCCTTGTGGACTTCCTGGCGCTGCCCTTCCCGTGGCCCCCGCT 420
Qy 1061 GluValIleLeuAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
Db 421 GAGGTGATCAAGGCTCCCGCATGCCCGGACCCCTCAGCTTCTCTACGCTCCACT 480
Qy 1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
Db 481 GGTCAACCACTGCTGGCTCCATGACACTGCTGGCGCTGCTCTGCGCGCCACCC 540
Qy 1101 ThrIleSerAsnProProProLeuIleSerSerAlaLeuHisProSerValLeuGluArg 1120
Db 541 ACCATCTCAACCCGCTCCCTCATCTCTCTGCAAGACACCCAGCGTCTCTGAGAGG 600
Qy 1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
Db 601 CAATAGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCAGTCCCGTACTCAGAGCAT 660
Qy 1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
Db 661 GCCAAGGCCCGGTGGGCTGTCAACCATGGGCTGGCCCTGCCCATGGACCCCAAAAG 720
Qy 1161 LeuAlaProPheSerGlyValLysGlnGlnLeuSerProArgGlyGlnAlaGlyPro 1180
Db 721 CTGGCACCTTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGCTGGGCCA 780
Qy 1181 ProGluSerLeuGlyValProThrAlaGlnGlnAlaSerValLeuArgGlyThrAlaLeu 1200
Db 781 CGGGAGAG-CTGGGGGTGCCACA-GCCCAAGNAGGCTCGTGCTGAGAGGACAGCTCTG 838
Qy 1201 GlySerValProGlyGlySerIleThrLysGlyIleProSer---ThrArgValProSer 1219
Db 839 GGTCTAGNTTCCGGCGGAGCATCAACAAAGGCAATCCCGCAGCACACGGGGTCCCTCG 898
Qy 1220 AspSer-AlaIleThrTyrArg-GlySerIleThrHisGly---ThrProAlaAspValL 1238
Db 899 GACACGNCATCATACCGGGGTTCATCACCACCGNNACGCCCGAGCTGGACGTCC 958
Qy 1238 euTyrLysGlyThr---IleThrArgIle---IleGlyGluAspSerProSerArgLeuA 1256
Db 959 TGGG-CAAGGNCACCCATCACCAGNATCATCGGGCGAGGACAGNCCGAGTGCCTTG 1017
Qy 1256 spArg-----GlyArgGluAspSerLeuProLys 1265
Db 1018 GAAACNCCCGCGGGAAGAACAGCCCTGGGCCCAAA 1055

RESULT 15
BM423558
LOCUS
DEFINITION AGENCOURT 6399411 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5516745
5', mRNA sequence.
ACCESSION BM423558
VERSION BM423558.1 GI:18391770
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2017 row: n column: 10
High quality sequence stop: 667.

```

FEATURES

source

Location/Qualifiers

```

1. .1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5516745"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

ORIGIN

Alignment Scores:

```

Pred. No.: 6.15e-52 Length: 1051
Score: 1479.50 Matches: 305
Percent Similarity: 78.95% Conservative: 10
Best Local Similarity: 76.44% Mismatches: 24
Query Match: 11.20% Indels: 60
DB: 12 Gaps: 4

```

US-09-522-753-5 (1-2517) x BM423558 (1-1051)

```

Qy 1457 LeuLysTyrAspThrGlyAlaSerThr-Thr-GlySerLysLysHisAspValArgSerLe 1476
Db 2 CTCAAGTACGACTCCGGCGCTCCACCACTGGGTCCAAAAGACGACGCTACGCTCCCT 61
Qy 1476 uIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAl 1496
Db 62 CATCGGACGCCCCGGCGGAGCTTCCACCCTGCGCTGGATGATGGCGGACGC 121
Qy 1496 aArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSe 1516
Db 122 CCGGGCATGGACGTGCTGTCTAGGAGGAGCTGAAGACCGCCGAGGACGCCAG 181
Qy 1516 rSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysPr 1536
Db 182 CAGCTCGGGGGCTCCATTGGCGCGCGCGCTCATTTGCTGCTGAGCTGGGCAAGCC 241
Qy 1536 oArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProAr 1556
Db 242 CGGCGACAGCCCCCTAACTATAGAGGACCCAGGGGACCCCTTTCGGCGCCACCTCCCA 301
Qy 1556 gGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSe 1576
Db 302 AGTTTCGGCGGTGACACCGGGGAGCCACCGCGGCTTGAGGAGGCGACCTTTCGTC 361
Qy 1576 rSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluLeuAlaLysSerPr 1596
Db 362 CAGCAAGGCATCCCGAGACCGAAAGCTGACGTCAGCGCTCGTGAGATCGCAAGTCCCC 421
Qy 1596 oHisSerThrValProGluHisHisProHisProLysSerProTyrGluHisLeuLeuAr 1616
Db 422 GCACAGACCGGTGCCCGGAGCACCCACACCCCATCTCGCCCTATGACGACCTTCGTC 481
Qy 1616 gGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSe 1636
Db 482 GGGGTGAGTGGCGTGGACCTGTATCGAGGCACATCCCCCTGGCTTCGACCCACCTC 541
Qy 1636 rIleProArgGlyIleProLeuAspAlaAlaAlaTyrTyrLeuProArgHisLeuAl 1656
Db 542 CATACCCGGGCGCATCCTCTGGAGCGAGCGCTGCCTACTACTCTGCGCCCGACACCTGG 601
Qy 1656 aProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspTh 1676
Db 602 CCCCACCCCACTTACCGGACACCTGTATCCCAACCCCTCATCTCCGCGGCTACCCGACAC 661

```



```
Db 606 -----|||||-----GTTGGCGCGGGAAGGCAAGGCTCTC 631
Qy 2406 rGlyArgProSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspAr 2426
Db 632 TGGCAGACCCAGCAGCCGANNAGCCAGTCCCCCGCCCGGCGCTGGCATCTGGGAGCG 691
Qy 2426 gProProSerValSerValHisSerGluGlyAspCysAsnArgArgThrProLeuTh 2446
Db 692 GCCACCCCTCTGTCTCTCAGTGCACTCGGAGGAGACTGCAACCCCGCGGAGCGCTCAC 751
Qy 2446 rAsnArgValTrpGluAspArgProSerSerLaglySerThrProPheProTyAsnPr 2466
Db 752 CAACCGCGTGTGGAGGACAGCCCTCCGTCCAGGTTCCAGCCCAITCCCTTACAACCC 811
Qy 2466 oleuilleMetArgLeuGlnAlaGlyValMetAlaSerProProPro-GlyLeuProA 2486
Db 812 CCTGATCATGGCGTCGAGCGCGGTGTCTATGGCTTCCCCACCCCGGCGCTCCCG 871
Qy 2486 laGlySer-GlyProLeuAlaGlyProHisAla---TrpAspGluGluProLysPro 2504
Db 872 CGGGCAGCGGGCCCTCGTGGCCCCCACCACGNCCTGGGAGCGGAGCCCAAGCCA 931
Qy 2505 Leu 2505
Db 932 CTG 934

BUES7144 846 bp mRNA linear EST 16-SEP-2002
AGNCOURT 10253238 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6584961 5', mRNA sequence.
BUES7144
BUES7144.1 GI:22907440
EST: Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 846)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L12M2796 row: c column: 09
High quality sequence stop: 719.
FEATURES
Location/Qualifiers
1..846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6584961"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/note="organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 9.76e-51 Length: 846
Score: 1448.00 Matches: 274
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.96% Indels: 0
DB: 13 Gaps: 0
US-09-522-753-5 (1-2517) x BUES7144 (1-846)
Qy 2244 LeuLeuTyArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGly 2263
Db 2 2 GTCTGTATACCGGGATGGGGAACAGACGAGGCCACAGAGTGGGTCTCAAGTCTCCAGGC 61
Qy 2264 AsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetVal 2283
Db 62 AACACCCAGCCAGCGCCAGCCTCTTTCAGCAAGCTGACCGAGAGCACTCCGCCATGTGC 121
Qy 2284 LysSerLysGlnGluLeuAsnLysLysLeuAsnThrHisAsnArgAsnGluProGlu 2303
Db 122 AAGTCCAGAAAGCAAGAGATCAACAAGAAGCTGAACACCCACACCGGAATGAGCCTGAA 181
Qy 2304 TyrAsnLysSerGlnProGlyThrGluLeuPheAsnMetProAlaLalleThrGlyThrGly 2323
Db 182 TACAATATCAGCCAGCTGGGAGCGAGATCTTCAATATGCCCGCATCACCAGAACAGGC 241
Qy 2324 LeuMetThrTyArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGlu 2343
Db 242 CTTATGACCTATAGAGCCAGCGGTGCAGGAACATGTCAGCACCAACATGGGGTGGAG 301
Qy 2344 AlaLalleArgLysAlaLeuMetGlyLysTyArgGlnTrpGluGluSerProProLeu 2363
Db 302 GCCATAATTAGAAAGGCATCATGGTAAATATGACCAAGTGGGAAGAGTCCCGCGCGTCTC 361
Qy 2364 SerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProLe 2383
Db 362 AGCCGCCAATGCTTTTAAACCTCTGAATGCCAGTGCCAGCTGCCCGCTGCTATGCCCAT 421
Qy 2384 ThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAla 2403
Db 422 ACCGCTGCTGACGAGCGAGTGACACACATCTACCTCCAGAGTGGCGGGGAGGCC 481
Qy 2404 LysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSer 2423
Db 482 AAGGTCTCTGGCAGACCCAGCAGCGAAAGTCCCGCGCGCGCTGGCATCT 541
Qy 2424 GlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsnArgThr 2443
Db 542 GGGAGCCGCGCCACCTCTGTCTCTCAGTGCACTCGGAGGAGAGCTGCAACCGCGGAGC 601
Qy 2444 ProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPhePro 2463
Db 602 CCGTCCACCAACCGCGTGTGGAGAGCAGGCCCTCTGTCGCGAGGTTCACGCCATTCGCC 661
Qy 2464 TyrAsnProLeuLeuMetArgLeuGlnAlaGlyValMetAlaSerProProProGly 2483
Db 662 TACAACCCCTGTATCATGCGGCTGCAGCGGGTGTGTCATGGTTCCTCCACCCCGGAGC 721
Qy 2484 LeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLys 2503
Db 722 CTCCCGCGGCGAGCGGCGCCCTCGCTGGCCCCCACCACCGCTGGGAGGAGGCCAAG 781
Qy 2504 ProLeuLeuCysSerGlnTrpGluThrLeuSerAspSerGlu 2517
Db 782 CCACTGCTCTCTCGCAGTAGCAGACACTCTCCGACAGCGAG 823
RESULT 18
BO690869 984 bp mRNA linear EST 15-JUL-2002
LOCUS BO690869
DEFINITION AGNCOURT_8034173 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6208241
5', mRNA sequence.
ACCESSION BO690869
VERSION BO690869.1 GI:21816185
```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 984)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCN2365 row: j column: 18
High quality sequence stop: 671.
Location/Qualifiers
1..984
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6208241"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1,956-50 Length: 984
Score: 1442.50 Matches: 296
Percent Similarity: 93.42% Conservative: 2
Best Local Similarity: 92.79% Mismatches: 12
Query Match: 10.92% Indels: 9
DB: 13 Gaps: 1

US-09-522-753-5 (1-2517) x BQ690869 (1-984)

QY 1334 AlaileProGluArgHisSerProHisHisLeuLysGluGlnHisHisileArgGly 1353
DB 1 GCCATCCCGCGGAGGACACAGCCCCCACCACCTCAAGAGACGACACATCCGCGG 60

QY 1354 SerileThrGlnGlyleProArgSerTyrValGluAlaGlnGluAspTyrLeuArg 1373
DB 61 TCATCACACAGGGATCCCTCGTCTACGTGGAGGACAGGAGACTACCTGGGTGG 120

QY 1374 GluAlaLysLeuLysArgGluGlyThrProProProProProSerArgAspLeu 1393
DB 121 GAGGCCAAGCTCTAAAGCGGGAGGACGCGCTCCGCGCCCGCCCTCACGGACCTG 180

QY 1394 ThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGly 1413
DB 181 ACCGAGCCCTACAGACGACGAGGCGCTGGCGCCCTGAACTGAGCGCGCCCATGGG 240

QY 1414 LeuValAlaThrValLysGluAlaGlyArgSerileHisGluLleProArgGluLeu 1433
DB 241 CTGGTGCCACCGGTGAAGAGGCGGCGCTCCATCATGAGATCCCGCGGAGGAGCTG 300

QY 1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerileThrGln 1453
DB 301 CGGCACACGCGCGAGCTGCGCCCTGGCGCCCGCGCGCGCTCAAGAGGAGGCTCCATCAGCGAG 360

QY 1454 GlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
DB 361 GGCACCCCGCTCAAGTACGACACCGCGCGCTCCACTGGCTCCAAAAGCACACGTA 420

QY 1474 ArgSerLeuLleGlySerProGlyArgThrPheProValHisProLeuAspValMet 1493
DB 421 CGCTCCCTCATCGGACGCGCGCGCGCGGAGTTCACCCCGCTGACCGCGCTGGATGATG 480

QY 1494 AlaAspAlaArgAlaLeuGluAlaGlyTyrGluSerLeuLysSerArgProGly 1513
DB 481 GCGAGCGCGCGCGCGCTGGAACGTGCTACGAGGAGAGCGCTGAAGAGCGCGCGAGG 540

QY 1514 ThrAlaSer-SerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLe 1533
DB 541 ACCGCGAGCGAGCTCGGGGGGCTCCATTGCGCGCGCGCGCTATTGTGCTCGAGCT 600

QY 1533 uGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHi 1553
DB 601 GGGTAAGCGCGCGCGAGAGCCCTGACCTATGAGGACACGCGGCGACCCCTTTGCCGCGCA 660

QY 1553 sLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGlySe 1573
DB 661 CTTCACGAGGTTTCGCGCGTGACACGCGGAGGCGCCACGCGCGCTCGAGGGCGAG 720

QY 1573 rLeuSerSerSerLysAlaSerGln-AspArgLysLeuThrSerThrProArgGluLea 1593
DB 721 CTTTCGTCCAGCAGGCTCCAGGAGCGGAAAGCTGACGTGCGCGCTCTCGAGATCG 780

QY 1593 lalysSerProHisSerThrValProGluHisHisProHis-ProileSer-ProTyrGI 1612
DB 781 CCAAGGTCCGCGCAGACCGCTGCCCGGAGCACACACCCCATCTCGCCCTATGA 840

QY 1612 uHisLeuLeuArgGlyVal---SerGlyValAsp-LeuTyrArgSerHis-IleProLeu 1630
DB 841 GCACCTGGCTCCGCGGCGGTAAATGGCTGGAACCTGTATCCACCACCATCCCGCTG 900

QY 1631 -AlaPheAspProThrSerIle-ProArgGlyIleProLeu 1643
DB 901 GGCCTTCAAAACCCACCTCCATACCCCGCGGCGCTCCTCTG 941

RESULT 19
BM802749 1159 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6459836 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575037
DEFINITION 5', mRNA sequence.
ACCESSION BM802749
VERSION BM802749.1 GI:19119572
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1159)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCN12324 row: k column: 06
High quality sequence stop: 652.
Location/Qualifiers
1..1159
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5575037"

FEATURES
source

/tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 2,828-50 Length: 1159
 Score: 1440.50 Matches: 291
 Percent Similarity: 87.57% Conservative: 12
 Best Local Similarity: 84.10% Mismatches: 20
 Query Match: 10.90% Indels: 23
 DB: 12 Gaps: 2

US-09-522-753-5 (1-2517) x BM802749 (1-1159)

Qy 1541 LeuThrTyArgHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
 |||||
 Db 2 CTGACCTATGAGGACACGGGCGACCTTTGCCGGCCACCTCCACGAGGTTCCGCCGTG 61
 |||||

Qy 1561 ThrMetArgGluProThrProArgLeuGlnGlySerLeuSerSerSerLeuAlaSer 1580
 |||||
 Db 62 ACCACGGCGAGCCACCGCGCGCTGCGAGGAGGCGAGCTTTCGTCAGCAAGGCATCC 121
 |||||

Qy 1581 GlnAspArgLysLeuThrSerThrProArgGluLeuAlaLysSerProHisSerThrVal 1600
 |||||
 Db 122 CAGGACCGAAGAGTGCATCGACGCTCGTGAGATCGCCAGATGCCCGCCACGACCGGTG 181
 |||||

Qy 1601 ProGluHisHisProHisProLysProThrTyArgHisLeuLeuArgGlyValSerGly 1620
 |||||
 Db 182 CCGGAGCACCCACCCACCTATCGCCCTATGAGCACCTGCTTCGGGGGCTGAGTGCC 241
 |||||

Qy 1621 ValAspLeuTyArgSerHisLeuProLeuAlaPheAspProThrSerIleProArgGly 1640
 |||||
 Db 242 GTGGACCTGTATCGCAGCCACATCCCTCGCTTCGACCCACCTCCATACCCCGCGGC 301
 |||||

Qy 1641 IleProLeuAspAlaAlaAlaTyTyTyLeuProArgHisLeuAlaProAsnProThr 1660
 |||||
 Db 302 ATCCCTCTGGAGCGAGCGCTGCTTACTCTGCTGCGCCGACACCTGCGCCCGCCACCCACCC 361
 |||||

Qy 1661 TyrProHisLeuTyProProTyLeuLeuArgGlyTyTyProAspThrAlaAlaLeuGlu 1680
 |||||
 Db 362 TACCGCACCTGTATCCACCTTACTCTATCGCGCTACCCGACACCGCGCGCTGAG 421
 |||||

Qy 1681 AsnArgGlnThrIleIleAsnAspTyTyTyLeuSerGlnGlnMetHisAsnThrAla 1700
 |||||
 Db 422 AACGGGACAGCATCATCAATGACTTACATCACCTCGCAGCAGATGCACCAACACAGCGCC 481
 |||||

Qy 1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
 |||||
 Db 482 ACCGCCATGCGCCAGCAGCTGATATGCTGAGGGCGCTCTCGCCCGCGAGTCTCGCTG 541
 |||||

Qy 1721 AlaLeuAsnTyAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
 |||||
 Db 542 GCACTCAACTAGCTGCGGGGCCCCGAGGACATCATGACCTGTGCCAAGTGCCACACCTG 601
 |||||

Qy 1741 ProValLeuValProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTy 1760
 |||||
 Db 602 CCGTGCTCGTGCGCCCGCAGACACCGACCCCGCAGCCACCGCATGGACCGCTTGCCTAC 661
 |||||

Qy 1761 LeuProThrAlaProGln-ProPheSerSerArgHisSerSerSerProLeuSerProGln 1780
 |||||
 Db 662 CTCCCCACCGCGCCCGCCCGCTTCAGCAGCGCGCCACAGAGCTCCCCACTCTCCCCAGG 721
 |||||

Qy 1780 yGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAsp 1800
 |||||
 Db 722 AGGTCCACACACTTGACAAACCAACCAACACACGTCCTCTCCGACCGGAGCGAAGCCG 781
 |||||

Qy 1800 gAspArgGlu-ArgAspArgAspArg-GluArgGluLysSerIleLeuThr-SerThr-Th 1819
 |||||

Db 782 GAATCGAGACGCGGACCGGATCGGGACCGGAAAGTCCTTCTTCCAGCTCCACCAAC 841
 |||||
 Qy 1819 rThrValGluHisAlaProLysLeuTrp-ArgProGly-ThrGluGlnSerSerGlySerSer 1838
 |||||
 Db 842 AACGGTGGAGCCCGGACCTCTCGGAACCTGGTAACCGAAGAAATAACCGGACGAA 901
 |||||
 Qy 1839 GlySerSerGlyGlyGlyGlySer-----SerSerArgPro 1851
 |||||
 Db 902 GGGAACAACAACGAGGGGGTGGTACACCTCTACAGAATACATACCTACGCTGCTGCT 961
 |||||

Qy 1852 -----AlaSerHisHisAlaHisGlnHisSer 1861
 |||||

Db 962 ATTGCTGTAGACGACGACTGGATAGTATTCGTCACATACCTCGTAATTATATCATAGT 1021
 |||||

Qy 1862 ProLys 1863
 |||||

Db 1022 CCGATC 1027

RESULT 20
 BM910785 1022 bp mRNA linear EST 12-MAR-2002
 BM910785
 LOCUS AGENCOURT_6617197 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5454337
 DEFINITION 5', mRNA sequence.
 ACCESSION BM910785
 VERSION BM910785.1 GI:19361164
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1022)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1949 row: f column: 02
 High quality sequence start: 12
 High quality sequence stop: 554.

FEATURES
 Location/Qualifiers
 1..1022
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_image="5454337"
 /tissue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_98"
 /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 1,74e-49 Length: 1022
 Score: 1420.00 Matches: 289
 Percent Similarity: 86.74% Conservative: 12
 Best Local Similarity: 83.29% Mismatches: 26
 Query Match: 10.75% Indels: 21
 DB: 12 Gaps: 4

993 CTCTGCCAAACCCCC 1007

BM558844 1030 bp mRNA linear EST 20-FEB-2002
 AGENCOURT 6592838 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5476852
 5', mRNA sequence.

BM558844
 BM558844.1 GI:18801929
 EST.

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/.
 1 (bases 1 to 1030)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM1993 row: p column: 05
 High quality sequence stop: 618.
 Location/Qualifiers
 1. .1030
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5476852"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 41"
 /notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 6.25e-48 Length: 1030
 Score: 1382.50 Matches: 288
 Percent Similarity: 85.76% Conservative: 7
 Best Local Similarity: 83.72% Mismatches: 29
 Query Match: 10.46% Indels: 21
 DB: 12 Gaps: 5

US-09-522-753-5 (1-2517) x BM558844 (1-1030)

Qy 1457 LeuIysTyrAspThrGlyAlaSerThrThr-GlySerIlysIyHispValArgSerLe 1476
 Db 2 CTCAGTACGACACCGCGCGGTCTCCACCACTGGGGTCCAAAAGACGACGACGACGTCCTCCCT 61

Qy 1476 uIlcIySerProGlyArgThrPheProValHisProLeuAspValMetAlaAspAl 1499
 Db 62 CATCGGAGACCCCGCGCGAGGTTCACCCGTCGACCCCGTGGATGTGATGGCCGACGC 121

Qy 1496 aArgAlaLeuGluArgAlaCysTyrGluGlySerLeuIysSerArgProGlyThrAlaSe 1511
 Db 122 CCGGGCACTGGAACTGCTCTACGAGGAGAGCGCTGAAGAGCGCGCCAGGACCGCCAG 181

Qy 1516 rSerSerGlyCysSerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysPr 1533
 Db 182 CAGCTCGGGGGGCTCCATTGCGCGCGGCCCGCCCGGTCAATTGTCTGAGCTGGGCAAGCC 241

1536 oArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProAr 1556
 242 GCGGAGAGCCCTTAACCTATGAGAGCACCGGGGACCCCTTTCGGCGACCTCCACG 301
 1556 gGlySerProValThrMetArgGluProThrProArgLeuGlnGlySerLeuSerSe 1576
 302 AGTTTCGGCCGCGGACACCGGGGAGCCACCGCGGCTGCGAGGAGGAGCGCTTTCGTC 361
 1576 rSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluLeuAlaLysSerPr 1596
 362 CAGCAGGATCTCCAGAGCCGAAAGCTGACGTCGAGCGCTCGTAGATCGCCAGTCCCC 421
 1596 oHisSerThrValProGluHisHisProHisProLysSerProTyrGluHisLeuLeuAr 1616
 422 GCACAGCACCGTCCGCGAGCACACCCACACCCCATCTCGCCCTATGACACCTGCTTCG 481
 1616 gGlyValSerGlyValAspLeuTyrArgSerHisLeuProLeuAlaPheAspProThrSe 1636
 482 GGGCGTGAGTGGCGTGGACCTGTATCGCAGGCACATCCCTCGGCTTCGACCCACCTC 541
 1636 rIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAl 1656
 542 CATACCCCGGCGATCCCTCTGGAGCGAGCGCTGCTACTACCTCCCGCGACACCTGGC 601
 1656 aProAsnProThrTyrProHisLeuTyrProProTyrLeuLeuAlaArgGlyTyrProAspTh 1676
 602 CCCCACCCACCTACCCGACCTGTATCCACCCCTACCTATCCGCGCTACCCCGACAC 661
 1676 rAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHi 1696
 662 GCGCGCTGGAGAACCGGACACCATCATCAATGACTACATCCCTCCGAGCAGATGCA 721
 1696 sHisAsnThrAlaThrAlaMetAlaGlnAlaAspMet-LeuArgGlyLeuSerProA 1716
 722 CCACACCGGCGCCACCGCATGGCCAGCAGCTGATATGTTGAGGGGCTCTCGCCCC 781
 1716 rg-GluSerSerLeuAlaLeuAsnTyrAla-----AlaGlyProArgGlyIleLeuAsp 1733
 782 GCGGAGACCTCGCTGGCACTCAACTACGCTGGCGCGTCCCGCCAGGAGGTC----- 835
 1734 LeuSerGlnValProHisLeu-----ProValLeuValProProThrProGly 1749
 836 -----CCACCCCACTTGACAAAACCAACCCCT--CTCCTTCCCGCCCTCCCGACG 884
 1750 ThrProAlaThrAlaMetAspArgLeuAlaTyrLeuPro-ThrAlaProGlnPro- 1767
 885 GGGGAGCGGAACCCGGGATTCACACAGGGCAACCGGCGATCGCGGAGCGGGGACA 944
 1768 ----PheSerSerArgHisSerSerSerProLeuSerProGly-----GlyProThrHi 1784
 945 CAATATTTCTTCGTTCCATGCTTCCCGCCAGATCGGTGGCCAGCCCGCGCGCCCA 1004
 1784 sLeu 1785
 1005 CCTC 1008

RESULT 22
 BM915686
 LOCUS
 DEFINITION AGENCOURT_6639688 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5481927
 5', mRNA sequence.
 ACCESSION BM915686
 VERSION BM915686.1 GI:19366065
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1010).
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: LNCM2007 row: c column: 16
 High quality sequence stop: 661.

FEATURES

source

1..1010
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5481927"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 8,95e-48 Length: 1010
 Score: 1378.50 Matches: 281
 Percent Similarity: 85.50% Conservative: 8
 Best Local Similarity: 83.14% Mismatches: 31
 Query Match: 10.43% Indels: 19
 DB: 12 Gaps: 3

US-09-522-753-5 (1-2517) x BM915686 (1-1010)

QY 1457 LeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeu 1476
 Db 1 CTCAGTAGACGACACCGGCGCTCCACCACTGGCTCCAAAAAGACGACGTACGCTCCCTC 60
 QY 1477 IleGlySerProGlyArgThrPheProValHisProLeuAspValMetAlaAspAla 1496
 Db 61 ATCGCAGCCCGCGCGGACGTTCCACCCGTGCACCCGCTGGATGTATGGCCGACGCC 120
 QY 1497 ArgAlaLeuGluArgAlaCysTyrGluSerLeuLysSerArgProGlyThrAlaSer 1516
 Db 121 CGGGCACTGGAACTGCTCTACGAGGAGAGCTGAAGAGCGCGGACGAGGACCGCCAGC 180
 QY 1517 SerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysPro 1536
 Db 181 AGCTCGGGGGCTCATTTGCGCGCGGCGCCCGGCTCATTTGTGCTGTAGCTGGCAAGCG 240
 QY 1537 ArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArg 1556
 Db 241 CGGCAGAGCCCTTAACCTATGAGGACACCGGGGACCCCTTTGCCGGCCACTCCACGA 300
 QY 1557 GlySerProValThrMetArgGluProThrProArgLeuGlnGlySerLeuSerSer 1576
 Db 301 GGTTCGCGCTGACACCGCGGAGCCACCGCGCGCTGCAGGAGGGGCGAGCTTTCGTC 360
 QY 1577 SerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluLeuAlaLysSerPro 1596
 Db 361 AGCAAGGCATCCAGAGACCGAAAGCTACGTCGACGCTCGTAGATCGCAAGTCCCGC 420
 QY 1597 HisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArg 1616
 Db 421 CACAGACCGGCGCGGAGCACACCCACCCCATCTCGCCCTATGAGACACTGCTTCGG 480
 QY 1617 GlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSer 1636

Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores: 1.19e-47 Length: 887
Pred. No.: 1374.00 Matches: 265
Score: 1374.00 Conservative: 0
Percent Similarity: 85.21% Mismatches: 0
Best Local Similarity: 85.21% Indels: 46
Query Match: 10.40% Gaps: 1
DB: 13

US-09-522-753-5 (1-2517) x BU542258 (1-887)

QY 2207 ProGluProAsnLysThrSerValLeuGlyGlyGluAspGlyLleGluProValSer 2226
Db 1 CCAGAGCCAAACAAGACGTCGGTCTTGGTGGTGGAGGACGGTATTGAACCTGTGTC 60
QY 2227 ProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyr 2246
Db 61 CCACCGGAGGGCATGACGGAGCCAGGCACTCCCGAGTGTGTGTACCGCTGTGTAC 120
QY 2247 ArgAspGlyGluInThrGluProSerArgMetGlySerLysSerProGlyAsnThrSer 2266
Db 121 CGGGATGGGMAACAGACGGAGCCAGCAGATGGGTCCCAAGTCTCCAGGCACACACG 180
QY 2267 GlnProProAlaPheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLys 2286
Db 181 CAGCCGCCAGCCTTCTCAGCAAGTCGACGAGCACTCCGCCATGTCAGTCCAGTCCAG 240
QY 2287 LysGlnGluLleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIle 2306
Db 241 AAGCAGAGATCAACAAGAGCTGAACACCCACACCGGAATGAGCCTGAATACATATC 300
QY 2307 SerGlnProGlyThrGluLlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThr 2326
Db 301 AGCCAGCTGGGAGGAGATCTTCAATATGCCCGCCATCAGGAGACAGGCTTATGACC 360
QY 2327 TyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIle 2346
Db 361 TATAGAACCCAGGCGGTGACGAAACATGCCAGACACCATGGGCTGGAGGCAATAAT 420
QY 2347 ArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluSerProProLeuSerAlaAsn 2366
Db 421 AGAAAGGCACCTCATG----- 435
QY 2367 AlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAla 2386
Db 435 ----- 435
QY 2387 AspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSer 2406
Db 436 -----GTTGGCGCGGGAAGGCAAGGCTCTCT 462
QY 2407 GlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArg 2426
Db 463 GGCAGACCCAGCAGCCGAAAGCCAAAGTCCCGGCCCGGGCCCTGGGATCTGGGGACCGG 522
QY 2427 ProProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThr 2446
Db 523 CCACCCCTCTGTCTCTCTAGTGCATCCGAGGAGACTGCAACCCCGGAGCGCCCTCAC 582
QY 2447 AsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnPro 2466
Db 583 AACCCGGTGTGGAGGACAGGCGCTCTGTCGCGAGGTTCCACGCCATTCCCCCTACACCCC 642
QY 2467 LeuLleMetArgLeuGlnAlaGlyValMetAlaSerProProProGlyLeuProAla 2486
Db 643 CTGATCATGCGGCTGCAGCGGGTGTCTATGGCTTCCCCACCCCGGCGCTTCCCGCG 702

Db 481 GGCCTGAGTGGCGTGGACCTGTATCGCAGCAGACATCCCCGTGGCTTCGACCCACCTCC 540
QY 1637 IleProArgGlyLeuProLeuAspAlaAlaAlaTyrTyrLeuProArgHisLeuAla 1656
Db 541 ATACCCCGCGGACATCCCTCTGGAGCAGCGCTGCTTACTACTGCTGCCGACACCTGGCC 600
QY 1657 ProAsnProThrTyrProHisLeuTyrProProTyrLeuLleArgGlyTyrProAspThr 1676
Db 601 CCCAACCCACCTACCCGACCTGTACCCACCTTACCTCATCGCGGCTACCCCGACAG 660
QY 1677 AlaAlaLeuGluAsnArgGlnThrLleAsnAspTyrLleThrSerGlnMetHis 1696
Db 661 GCGCGCTGGAGAACCGGACACCATCATCAATCACTACATCACTCGCAGCAGATGCAC 720
QY 1697 HisAsnThrAlaThrAlaMetAlaGlnArg-AlaAsp-MetLeuArgGlyLeuSerProA 1716
Db 721 CACAACGGGTCCCGCCCATGGCCAGCTGAGCTAATATTGCTGAGGGGCTCTCGGCC 780
QY 1716 TgGluSerSerLeuAlaLeuAsnTyrAlaAlaGly----ProArgGlyLleLleAspLeuS 1735
Db 781 GCGAGTCTCGTGGCACTCAACTACCTCGCGGGTCCCGCAGGAGGTCAAAACCTTG 840
QY 1735 erGlnValProHisLeuProValLeu-----ValProProThrPro---- 1748
Db 841 GCCAAA--CCAAACCGCGCGGTCTTCGCCCAACGGGCGAGTGGAGCCCGGAATCCAAACG 898
QY 1749 -----GlyThrProAlaThrAlaMetAspArgLeuAlar 1760
Db 899 CGTACCCGATCGGCGAGGGGAAAGGCACTGTTTACGTCGCCACCGGCGAGTCAAAAGGTGAAAC 958
QY 1760 yriLeuProThrAlaProGlnProPheSerSerArgHisSerSerPro 1776
Db 959 ATGGGCGCCCTCTGTTAAACCGTGAAACATGAGGAGACTACGGGTCCC 1008

RESULT 23
BU542258 887 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10253089 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574079
DEFINITION 5', mRNA sequence.
ACCESSION BU542258
VERSION BU542258.1 GI:22852741
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
NTH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2768 row: m column: 23
High quality sequence stop: 710.
Location/Qualifiers
1. .887
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574079"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming."

QY 2487 GlySerGlyProLeuAlaGlyProHisAlaIatpAspGluGluProLysProLeuLeu 2506
 Db 703 GGCAGCGGCGCCCTGGTGGCCGCCACCAACCGCTGGAGCAGGAGCGCAAGCACTGTCTC 762
 QY 2507 CysSerGlnTyrGluThrLeuSerAspSerGlu 2517
 Db 763 TGCTCGCAGTACGAGACACTCTCCGACAGCGAG 795

RESULT 24

BM472005 1067 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT_6464448 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5581660
 DEFINITION 5', mRNA sequence.

ACCESSION BM472005
 VERSION EST.
 KEYWORDS BM472005.1 GI:18521047
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1067)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12341 row: o column: 05
 High quality sequence stop: 701.

FEATURES

source

Location/Qualifiers
 1..1067
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5581660"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 67"
 /note="Organ: eye; Vector: pCMV-Sport6; Site 1: Not I;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 Kb. Library constructed by Life
 Technologies."

ORIGIN

Alignment Scores:
 Pred. No.: 2,596-47 Length: 1067
 Score: 1368.00 Matches: 289
 Percent Similarity: 82.83% Conservative: 10
 Best Local Similarity: 80.06% Mismatches: 38
 Query Match: 10.35% Indels: 24
 DB: 12 Gaps: 7

US-09-522-753-5 (1-2517) x BM472005 (1-1067)

QY 892 LysAlaGluLysLysGluGlySerGlyArgAlaThrAlaLysSerSerGlyAla 911
 Db 1 AAGGCAGAGAGAGAGAGCGGCGGAGCGCAGCGCCACACAGCAAGCTCGGGCGCC 60
 QY 912 ProGlnAspSerAspSerSerAlaThrCysSerAlaAspGluValAspGluAlaGly 931
 Db 61 CCCAGGACAGCGACTCCAGTGCACCTGCAGTGCAGAGCGAGTGCATGAGCGCGCGC 120
 QY 932 GlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuThrProThrGlyAspPro 951
 Db 121 GCGCAGAGAACCGCGCTGTGTGCCCAAGCGCCAGCCCTCTCACCCCGAGCGCGACCC 180

QY 952 ArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLeuLysGlnAlaAla 971
 Db 181 CGGGCAATGCTCTACCCCGAGAGCACTGGACCTGAAGCAGCTGAAGCAGCGGCT 240
 QY 972 AlaIleProProlIleGlnValThrLysValHisGluProProArgGluAspAlaPro 991
 Db 241 GCCATCCCCCATCCAGGTCAACAAAGTCCATGAGCCCCCGGGAGACGACGCTCC 300
 QY 992 ThrLysProAlaProProAlaProProProGlnAsnLeuGlnProGlnLysSerAspAla 1011
 Db 301 ACCAAGCAGCTCCCCAGCCCGCCACCGCAACAACTTCGACCGCGGAGCGCGCC 360
 QY 1012 ProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAlaProProAlaAsp 1031
 Db 361 CCTCAGCAGCTGGCAGCAGCGCCCGCGGCAAGAGCAGGAGCCCGCACCCCGCCGAC 420
 QY 1032 LysGlu-----AlaPheAlaAlaGluAlaGlnLysLeuPro 1043
 Db 421 AAGGAGGCGAGAGAGCCGTGTCTTCCACAGCCTTCGACGCGGAGCGCCAGAACTGGCT 480
 QY 1044 GlyAspProProCysTyrThrSerGlyLeuProPheProValProProArgGluValIle 1063
 Db 481 GGGGACCCCTCTGCTGACTTCGGCCCTGCTCTCCCGTGCCTCCCGCGGTGATC 540
 QY 1064 LysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProProGlyHisPro 1083
 Db 541 AAGCCTCCCGCATGCCCGGAGCCCTCAGCTCTCTCTAGCTCCACCTGGTCACCCA 600
 QY 1084 LeuProLeuGlyLeuHisAspThrAlaArgProValLeuProProArgProThrIleSer 1103
 Db 601 CTGCCCCCTGGGCTCCATGACACTGCCCGCGCCGCTCTGCGCGGCCACCCACCATCTCC 660
 QY 1104 AsnProProLeuLeuSerSerAlaLysHisProSerValLeuGluArgGlnIleG1 1123
 Db 661 AACCGCGCTCCCTCATCTCTCTGCCAAGCACCCAGCGTCTCTCAGAGGCAATAGG 720
 QY 1123 YAlaIleSerGln-GlyMetSerValGlnLeuHisValProTyrSerGluHisAlaLys- 1142
 Db 721 TGCCATCTCCCAAGGAATGTCGGTCCAGTCCAGTCCCGTCTCTCAGAGCAGCAAGGG 780
 QY 1143 AlaProValGlyProValThrMet---GlyLeuProLeuProMetAspProLysLysLeu 1161
 Db 781 GCCCGGFGGGCCCTGTCCACCATGGGGGCTGCCCCCTGCCCATGGAACCCCAAAAGCT 840
 QY 1162 --AlaProPheSerGlyValLys-GlnGluGlnLeuSerPro---ArgGlyGlnAla-- 1178
 Db 841 GGGACCCCTTCAACGGGAGTGAAACCAAGAACAGTGTCTCCACCGCGGGCGGCGCTG 900
 QY 1179 GlyProProGluSerLeu-----GlyValProThrAlaGlnGluAlaSerValLeuArg 1196
 Db 901 GGGCCACCCGAAAAAACCCCTGGGGGGGCGCCACACGCCCCCGAAAGCC-----GTCCCG 954
 QY 1197 GlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGly-----IlePro 1213
 Db 955 GGCCTGGAGAGAGGACAGCTTCCGGGGCTTCAATTTCCGGGGGGGGAAGACCTTCCGCC 1014
 QY 1214 SerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThr 1230
 Db 1015 CAAAGGGATTTTCCAGGAACACGCGGGGGGCCCCCTCCGAAAGAGCGCCACT 1065

RESULT 25

BX368971

LOCUS

DEFINITION BX368971 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL010YH07 5-PRIME, mRNA sequence.

ACCESSION BX368971

VERSION BX368971.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 868)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6282.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0BAE0292G02_AE02704_1&cluster=6282.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0BAE0292G02_AE02704_1. Location/Qualifiers		1. 868 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DL010YH07" /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED" /cell_line="RAMOS CELL LINE" /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		2. 38e-47 1366.50 92.81% 91.78% 10.34% 13		Length: Matches: Conservative: Mismatches: Indels: Gaps:		868 268 3 13 9 3													
FEATURES		source		1. 868		/organism="Homo sapiens"		/mol_type="mRNA"		/db_xref="taxon:9606"		/clone="CS0DL010YH07"		/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"		/cell_line="RAMOS CELL LINE"		/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"		/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
ALIGNMENT		Alignment Scores:		Pred. No.:		Score:		Percent Similarity:		Best Local Similarity:		Query Match:		DB:		US-09-522-753-5 (1-2517) x BX368971 (1-868)		1590 ArgGluLeuAlaLysSerProHisSerThrValProGluHisHisProHisProHisSer		1609	
Y		Y		Y		Y		Y		Y		Y		Y		Y		2 CGTGGGATGCACAGTCCCGAC-AGCACCGTGGCGGAGCACCACCCACCCATCTCG		60	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1610 ProTyrGluHisLeuAlaArgGlyValSerGlyValAspLeuTyrArgSerHisIlePro		1629	
Y		Y		Y		Y		Y		Y		Y		Y		Y		61 CCCTATGAGCACCTGCTCGGGCGGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCC		120	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1630 LeuAlaPheAspProThrSerIleProArgGlyLeuProLeuAspAla-AlaAlaAlaTy		1649	
Y		Y		Y		Y		Y		Y		Y		Y		Y		121 CTGGCTTGCACCCACCTCCATACCCCGGGGATCCCTCTGGACGCCGCGCTGCCTA		180	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1649 rTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrIle		1669	
Y		Y		Y		Y		Y		Y		Y		Y		Y		181 CTACCTGCCCGACACCTGGCCCCCAACCCACCTACCGCACCTGTACCCACCTACCT		240	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1669 uileArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleLeuAsnAspTy		1689	
Y		Y		Y		Y		Y		Y		Y		Y		Y		241 CATCCGGCGGTACCCCGACACCGCGCGGTGGAGAACCGGCAGACCATCATCAATGACTA		300	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1689 rIleThrSerGlnGlnMetHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMe		1709	
Y		Y		Y		Y		Y		Y		Y		Y		Y		301 CATACCTCCGACGATGACCAACACCGCGGCACCGCATATGCCAGGAGCTGATAT		360	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1709 tLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProAr		1729	
Y		Y		Y		Y		Y		Y		Y		Y		Y		361 GCTGAGGGGCTCTCGCCCCCGAGTCTCTCGCTGGCACTCAACTAGCGTGGCGTCCCCG		420	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1729 gGlyIleLeuAspLeuSerGlnValProHisLeuProValLeuValProThrProGl		1749	
Y		Y		Y		Y		Y		Y		Y		Y		Y		421 AGGCATCATGACCTGTCCCAAGTGCACACCTGCTGTGCTGCTGCTGCTGCTGCTGCTG		480	
Y		Y		Y		Y		Y		Y		Y		Y		Y		FEATURES		source	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1749		1769	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1789		1809	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1809		1829	
Y		Y		Y		Y		Y		Y		Y		Y		Y		601		660	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1809		1847	
Y		Y		Y		Y		Y		Y		Y		Y		Y		721		780	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1847		1864	
Y		Y		Y		Y		Y		Y		Y		Y		Y		781		840	
Y		Y		Y		Y		Y		Y		Y		Y		Y		1864		1875	
Y		Y		Y		Y		Y		Y		Y		Y		Y		841		868	
Y		Y		Y		Y		Y		Y		Y		Y		Y		923 bp		mRNA	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AGENCOURT_7939306 NIH_MCC_67 Homo sapiens cDNA clone		linear	
Y		Y		Y		Y		Y		Y		Y		Y		Y		5', mRNA sequence.		EST 04-SEP-2002	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		6172730	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		1	
Y		Y		Y		Y		Y		Y		Y		Y		Y		GI:22698387		EST.	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens (human)		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		rProArgThrGlnAspAlaLeuGlnGlnArgPro		1875	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AAGGCCCTCCACACAGAGACCA		868	
Y		Y		Y		Y		Y		Y		Y		Y		Y		RESULT 26		923 bp	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		mRNA	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AGENCOURT_7939306 NIH_MCC_67 Homo sapiens cDNA clone		linear	
Y		Y		Y		Y		Y		Y		Y		Y		Y		5', mRNA sequence.		EST 04-SEP-2002	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		6172730	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		1	
Y		Y		Y		Y		Y		Y		Y		Y		Y		GI:22698387		EST.	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens (human)		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		rProArgThrGlnAspAlaLeuGlnGlnArgPro		1875	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AAGGCCCTCCACACAGAGACCA		868	
Y		Y		Y		Y		Y		Y		Y		Y		Y		RESULT 26		923 bp	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		mRNA	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AGENCOURT_7939306 NIH_MCC_67 Homo sapiens cDNA clone		linear	
Y		Y		Y		Y		Y		Y		Y		Y		Y		5', mRNA sequence.		EST 04-SEP-2002	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		6172730	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		1	
Y		Y		Y		Y		Y		Y		Y		Y		Y		GI:22698387		EST.	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens (human)		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		rProArgThrGlnAspAlaLeuGlnGlnArgPro		1875	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AAGGCCCTCCACACAGAGACCA		868	
Y		Y		Y		Y		Y		Y		Y		Y		Y		RESULT 26		923 bp	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		mRNA	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AGENCOURT_7939306 NIH_MCC_67 Homo sapiens cDNA clone		linear	
Y		Y		Y		Y		Y		Y		Y		Y		Y		5', mRNA sequence.		EST 04-SEP-2002	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		6172730	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		1	
Y		Y		Y		Y		Y		Y		Y		Y		Y		GI:22698387		EST.	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens (human)		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		rProArgThrGlnAspAlaLeuGlnGlnArgPro		1875	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AAGGCCCTCCACACAGAGACCA		868	
Y		Y		Y		Y		Y		Y		Y		Y		Y		RESULT 26		923 bp	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		mRNA	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AGENCOURT_7939306 NIH_MCC_67 Homo sapiens cDNA clone		linear	
Y		Y		Y		Y		Y		Y		Y		Y		Y		5', mRNA sequence.		EST 04-SEP-2002	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		6172730	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		1	
Y		Y		Y		Y		Y		Y		Y		Y		Y		GI:22698387		EST.	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens (human)		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		rProArgThrGlnAspAlaLeuGlnGlnArgPro		1875	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AAGGCCCTCCACACAGAGACCA		868	
Y		Y		Y		Y		Y		Y		Y		Y		Y		RESULT 26		923 bp	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		mRNA	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AGENCOURT_7939306 NIH_MCC_67 Homo sapiens cDNA clone		linear	
Y		Y		Y		Y		Y		Y		Y		Y		Y		5', mRNA sequence.		EST 04-SEP-2002	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		6172730	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		1	
Y		Y		Y		Y		Y		Y		Y		Y		Y		GI:22698387		EST.	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens (human)		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		Homo sapiens		Homo sapiens	
Y		Y		Y		Y		Y		Y		Y		Y		Y		rProArgThrGlnAspAlaLeuGlnGlnArgPro		1875	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AAGGCCCTCCACACAGAGACCA		868	
Y		Y		Y		Y		Y		Y		Y		Y		Y		RESULT 26		923 bp	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		mRNA	
Y		Y		Y		Y		Y		Y		Y		Y		Y		AGENCOURT_7939306 NIH_MCC_67 Homo sapiens cDNA clone		linear	
Y		Y		Y		Y		Y		Y		Y		Y		Y		5', mRNA sequence.		EST 04-SEP-2002	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU184403		6172730	
Y		Y		Y		Y		Y		Y		Y		Y		Y		BU1			

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: NDAM055 row: j column: 05
High quality sequence stop: 649.
Location/Qualifiers

FEATURES

1..887
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30143932"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 164"
Note=Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Non-normalized full-length enriched library from pooled
mouse embryonic limb, maxilla and mandible, day 10.5 and
11.5 (size selected for the 0.5-1 kb fragments). Cloned
directionally, priming method: Oligo-dT. cDNA enrichment:
5' GACTAGTCTTAGATCGGAGCGCCGCTT 3'. Tissue contributed
by David Rowe. Library constructed by ResGen, Invitrogen
Corp."

ORIGIN

Alignment Scores:
Pred. No.: 3,69e-46 Length: 887
Score: 1338.00 Matches: 262
Percent Similarity: 89.00% Conservative: 13
Best Local Similarity: 84.79% Mismatches: 16
Query Match: 10.12% Indels: 18
DB: 14 Gaps: 5

US-09-522-753-5 (1-2517) x CA979881 (1-887)

311 CAGACGGAGCCAGCAGGATGGGCTCCAGTCTCCAGGCAACACAGCCAGCCGCGAGCC 370
2271 PhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysGlnGluIle 2290
371 TTCCTCAGCAAGCTGACGAGAGCACTCCGCCATGCTCAAGTCCAGGAAGCAAGAGATC 430
2291 AsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGly 2310
431 AACAGAGCTGAACACCCACCCAGCAATGAGCTGAATACATATATACAGCCAGCTGGG 490
2311 ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGln 2330
491 ACGGAGATCTTCAATATATCCCGCATCACCGGAACAGCCCTATACCTATAGAGCCAG 550
2331 AlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeu 2350
551 GCGGTGTCAGGACATGCCCGCACCAACATGGGCTGGAGGCCATAATAGAAAGGCACTC 610
2351 MetGlyLysTyrAspGlnTyrGluSerProProLeuSerAlaAsnAlaPheAsnPro 2370
611 ATGGGTAATATGACAGTGGGAGAGTCCCGCGGCTCAGCGCAATGCTTTTAAACCT 670
2371 LeuAsnAlaSerAla-SerLeuProAlaAlaMetProIleThrAlaAla-AspGlyArgS 2390
671 CTGAATGCCAGTGGCCGCTGCCCTGCTATGCCCATACCGTTGCTGACGAGCGGA 730
2390 exAspHisThrLeuThrSerProGlyGly--GlyGlyLysAlaLysValSerGlyArg-- 2408
731 GTGACACACACTCACTCCCTCCAGTGGGCGGCGGAAAGCCAGGGTCTCTGGGCAAC 790
2409 -ProSerArgLysAlaLysSerProAlaProGlyLeuAlaSer---GlyAspArgPr 2427
791 CCACAAACCGAAAGCCAAATTTCCCGCGCCCGGCTGGGAATCTGGGGGACCCGCC 850
2427 oProSerValSerValHisSerGluGlyAspCys-Asn-----ArgArgT 2443
851 CCCCTTTTGGTTCTCTCAGGGACCTCCCGAGCGAAACTTGTAACCCCTCGGGAAC 910
2443 hrProLeu---ThrAsnArgValTyr-----GluAspArgProSerSerAlaGlySerT 2460
911 GCCTCTCTCCCAACACTCTTTGTGGGCAAGAAACCCGCCCTTTTNTCCCGCGGCC 970
2460 hrProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProp 2480
971 TTCCTCCGCCCCATTTCCCTCCCTA-----ACACCCC 1000
2480 roProProGlyLeu 2484
1001 CCCCCCGGCTCTCTC 1014

RESULT 28
CA979881
LOCUS
DEFINITION
CA979881 887 bp mRNA linear EST 06-JAN-2003
AGENCOURT 11276590 NIH MGC 164 Mus musculus cDNA clone
IMAGE:30143932 5', mRNA sequence.
CA979881
VERSION 1 GI:27512535
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NTH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: ccapbs@mail.nih.gov
Tissue Procurement: Dr. David Rowe and Dr. Mina
cDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

1813	Qy	SerIleLeuThrSerThrThrValGluHisAlaProIleTrpArgProGlyThrGlu	1832
583	Db	TCCATCTCCAGCTACCACTACAGTGGAGCATGTACCACTCTGGAGACCTGTCGGAG	642
1833	Qy	GlnSerSerGlySerSerGlySerSerGlyGlyGlySerSerSerArgProAla	1852
643	Db	CAGAGCAGCGGGGCT-----GGGGCGAGCAGCCGCCCGGCC	678
1853	Qy	SerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGln	1872
679	Db	TCCACACC-----CACGAGCACTGCCCATCTCCCCCGGAGCCAGGACGCTTGCAG	732
1873	Qy	GlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPro	1892
733	Db	CAGAGCCCACTGTGCTGCACACACAGCATGAGGGCGTGTCACTCCGTGGHACCC	792
1893	Qy	SerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThr	1912
793	Db	GGCAGCGCCACCGTCTCTAGGTCCCCACTCACCTTTCGCTCTCCGCCCGAGTGCACA	852
1913	Qy	PheProPro---AlaThrHisCysPro	1920
853	Db	TTTCCAACTCTGGCCACCACCAACTCGCCC	879

RESULT 29					
BX390462					
LOCUS					
DEFINITION	BX390462	926 bp	mRNA	linear	EST 23-JUN-2003
	BX390462	Hom sapiens PLACENTA COT 25-NORMALIZED	Hom sapiens cDNA		
	clone CS0DI056YJ24	5-PRIME,	mRNA sequence.		

ACCESSION	BX390462
VERSION	BX390462.1
KEYWORDS	GI:32165291
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 926)
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 7437.f,
 Contact : Peng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/>
 Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAG0322G08_CS03039_1'.
 Location/Qualifiers
 1..926
 source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/gene="CSODI056YJ24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN	
Alignment Scores:	
Pred. No.:	5,566-46
Score:	1334.00
Percent Similarity:	89.61%
Best Local Similarity:	90.32%
Query Match:	10.09%
DB:	13
	Length: 926
	Matches: 276
	Conservative: 4
	Mismatches: 21
	Indels: 13
	Gaps: 1
US-09-522-753-5 (1-2517) x BX390462 (1-926)	

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLML at:
http://image.llnl.gov
Plate: LLCW2385 row: d column: 14
High quality sequence stop: 668.
FEATURES
Location/Qualifiers
1..875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6215773"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 6 44e-46 Length: 875
Score: 1332.00 Matches: 252
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.08% Indels: 0
DB: 13 Gaps: 0
US-09-522-753-5 (1-2517) x BQ711119 (1-875)
QY 2266 SerGlnProAlaPheSerLysLeuThrGluSerAsnSerAlaMetValLysSer 2285
Db 3 AGCCAGCCGCCAGCTTCTTCAGCAAGCTGACCGAGCACTCCGCCATGGTCAAGTCC 62
QY 2286 LysLysGlnGluLeuAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsn 2305
Db 63 AAGAGCAAGAGATCAACAGAGAGCTGAACACCCACACCGAAGTGGCTGATACAT 122
QY 2306 IleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMet 2325
Db 123 ATCAGCAGCCCTGGGAGCGAGATCTTCATATGCGCGCATCACCGAAGAGCGCTTATG 182
QY 2326 ThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIle 2345
Db 183 ACCTATAGAGCCAGCGGTGTCAGGAACATGCCAGACCAACATGGGGCTGGAGGCCATA 242
QY 2346 IleArgLysAlaLeuMetGlyLysTyrAspGlnTyrGluSerProProLeuSerAla 2365
Db 243 ATTAGAAGGCATCATGGTAAATATGACCATGGAGAGATCCCGCGCTCAGCGCC 302
QY 2366 AsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAla 2385
Db 303 AATGCTTTTAACTCTGAATGCCAGTGCAGCGCTGCTATGCTATGCCCATTAACCGCT 362
QY 2386 AlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysVal 2405
Db 363 GCTGACGGCGGAGTGACCAACACTCCTCCAGGTGGCGCGGAGGCCAAGTCTC 422

```

```

QY 2406 SerGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAsp 2425
Db 423 TCTGGCAGACCCAGCAGCCGAAAGCAAGTCCCGCGCCGCGCTGGCATCTGGGAC 482
QY 2426 ArgProProSerValSerSerValHisSerGluGlyAspCysAsnArgThrProLeu 2445
Db 483 CGGCCACCCCTCTGTCTCTCAGTCGAGGAGACTGCAACCGCCGAGCGGCTC 542
QY 2446 ThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsn 2465
Db 543 ACCAACCGCGGTGGGAGGAGCAGCCCTCTGTCGAGAGTTCACGCCATTCCTCCACAC 602
QY 2466 ProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProProProGlyLeuPro 2485
Db 603 CCCTCATCATCGCGCTGCAGCGGGTGTCTGTGCTTCCACCCCGCGGCTCCCTCC 662
QY 2486 AlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeu 2505
Db 663 GCGGCGAGCGGCGCCCTCGTGGCCCCCACCAGCGCTGGGAGCGAGGCCACCTG 722
QY 2506 LeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
Db 723 CTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 758

```

RESULT 31

BU224569

LOCUS

DEFINITION

603400211F1 CSQCHN23 Gallus gallus cDNA clone CHEST293f3 5', mRNA

sequence.

ACCESSION

BU224569

VERSION

BU224569.1 GI:25459595

KEYWORDS

EST.

SOURCE

Gallus gallus (chicken)

ORGANISM

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

REFERENCE

1 (bases 1 to 842)

AUTHORS

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

TITLE

A Comprehensive Collection of Chicken cDNAs

JOURNAL

Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE

22335534

PUBMED

12445392

COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..842

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hise"

/db_xref="taxon:9031"

/clone="CHEST293f3"

/dev_stage="22"

/lab_host="DH10B"

/clone_lib="CSQCHN23"

/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:

EcoRI; Site_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunted, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2


```
QY 2348 LysAlaLeuMetGlyLysTyrAspGlnTrpGluSerProLeuSerAlaAsnAla 2367
Db 405 -----
QY 2368 PheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaasp 2387
Db 405 -----
QY 2388 GlyArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGly 2407
Db 406 -----GTTGGGGGGGGAAGCCAGGCTCTTGGC 435
QY 2408 ArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgPro 2427
Db 436 AGACCCAGCAGCGAAGCCAAAGTCCCGCGCCCGGGCTGGCATCTGGGGACCGGCCA 495
QY 2428 ProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsn 2447
Db 496 CCTCTGTCTCTCAGTGACCTCGAGGGAGACTGCAACCGCGGACCGCGCTCACCAAC 555
QY 2448 ArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeu 2467
Db 556 CGCGTGTGGAGGACAGGCCCTCTGTCGGCAGGTTTCAGGCCATTCCCTACCAACCCCTG 615
QY 2468 IleMetArgLeuGlnAlaGlyValMetAlaSerProProProGlyLeuProAlaGly 2487
Db 616 ATCATGCGGCTGACGCGGGTGTATGGCTTCCCCACCCACCGCGGCTTCCCGCGGGC 675
QY 2488 SerGlyProLeuAlaGlyProHisAlaTrpAspGluProLysProLeuLeuCys 2507
Db 676 AGCGGGCCCTCGCTGGGCCCCCACCACGCTGGGACGAGGACCAAGCCACTGCTCTGC 735
QY 2508 SerGlnTyrGluThrLeuSerAspSerGlu 2517
Db 736 TCGCAGTACGACACTCTCCGACAGCGAG 765
```

```
RESULT 34
LOCUS BUI80236
DEFINITION AGENCOURT_8042921 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084555
5', mRNA sequence.
ACCESSION BUI80236
VERSION BUI80236.1 GI:22694220
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2316 row: i column: 04
High quality sequence stop: 622.
FEATURES
source
Location/Qualifiers
1..882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6084555"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
```

/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

```
Alignment Scores:
Pred. No.: 1.85e-45 Length: 882
Score: 1321.00 Matches: 266
Percent Similarity: 94.74% Conservative: 4
Best Local Similarity: 93.33% Mismatches: 9
Query Match: 10.00% Indels: 6
DB: 13 Gaps: 0

US-09-522-753-5 (1-2517) x BUI80236 (1-882)

QY 1334 AlalleProGluuArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGly 1353
Db 1 GCCATCCCGCGGAGCGACACAGCCCCCACCACCTCAAGAGAGACACACATCCCGGG 60
QY 1354 SerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArg 1373
Db 61 TCCATCACACAAGGATCCCTCGGTCTTACGTGGAGGCACAGGAGGACTTACTGCTCG 120
QY 1374 GluAlaLysLeuLysArgGlyThrProProProProProProProProProProPro 1393
Db 121 GAGGCCAAGCTCTTAAAGCGGAGGACGCGCTCGGCCCCACCGCCCTCAGGGACCTG 180
QY 1394 ThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGly 1413
Db 181 ACCGAGGCTTAAAGACGACGAGCCCTGGGCCCCCTGAAGCTGAAGCCCGCCCATGAG 240
QY 1414 LeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGlu 1433
Db 241 CTGGTGGCCACCGTGAAGGAGCGGCGCTCCATCATGAGATCCCGCGGAGGAGCTG 300
QY 1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGln 1453
Db 301 CGGCACACGCGCGAGCTGCCCTCGGCCCGCGCGCTCAAGGAGGCGCTCATCATCAG 360
QY 1454 GlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
Db 361 GGCACCGCGCTCAGTAGACACACCGCGCGCTCCACCTGGCTCCAAAAGACGACGAC 420
QY 1474 ArgSerLeuLysGlySerProGlyArgThrPheProProValHisProLeuAspValMet 1493
Db 421 CGCTCCCTCATCGGACGCGCGCGCGGAGCTTCCACCCCGTGACCCCGCTGGATGATG 480
QY 1494 AlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGly 1513
Db 481 GCCGACGCGCGGACCTGGAAACGTCCTGCTACGAGGAGAGCCTCAAGAGCGCGCAG 540
QY 1514 ThrAlaSer-SerSerGlyGlySerIleAlaArgValAlaProValIleValProGluLe 1533
Db 541 ACCGCCAGAGGCTCGGGGGGCTCATTTGCGCGCGCGCGCGCGCTCATTTGCTTGA 600
QY 1533 uGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAla-ProPheAlaGlyH 1553
Db 601 GGGTAAGCGCGGACAGGCCCCCTGACCTATGAGGACACCGCGGCGCCCTTTGCCG 660
QY 1553 isLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGlyLys 1573
Db 661 ACCTCCACGAAAGTTCGCGCGTGGACCGCGGAGCCACCGCGCTCGAGGAAGCA 720
QY 1573 erLeuSerSerSerLysAlaSer-GlnAspArgLysLeuThrSerThrProArgGluLutle 1592
Db 721 GCCTTTTCTCCCGAAGGGCTCCCGCAGGACCGAAGCTGACGTAACCCCTCGGGAGATC 780
QY 1593 AlalysSerProHis-SerThrValProGluHisHisProHis-ProIleSer-ProTyr 1611
```

```

Db      781  GCCAAGGCGCCGACAACTACCGTGCCTGCGGAGATCAACCAACCAATTTGCGCCCTAT 840
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      1612  Glu 1612
      |||
Db      841  GAA 843

RESULT 35
BX368972      943 bp      mRNA      linear      EST 08-MAY-2003
LOCUS      BX368972 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION      Homo sapiens cDNA clone CS0DL010H07 5-PRIME, mRNA sequence.
ACCESSION      BX368972
VERSION      BX368972.1 GI:30449776
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 943)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6282.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAB029ZG02_AE02704_2&cluster=6282.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0BAB029ZG02_AE02704_2.

FEATURES
Location/Qualifiers
1..943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL010YH07"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.: 4,49e-45 Length: 943
Score: 1312.50 Matches: 273
Percent Similarity: 80.87% Conservative: 6
Best Local Similarity: 79.13% Mismatches: 32
Query Match: 9.93% Indels: 36
DB: 13 Gaps: 5

US-09-522-753-5 (1-2517) x BX368972 (1-943)
Qy      1579  AlaSerGlnAepArgLysLeuThrSerThrProArgGluLeuAlaLysSerProHisSer 1598
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      3      GCATCCAGACCG-AGCTGACGTGACGCGCTGTCGAGGACGCCAAGTCCCGCACAGC 61
Qy      1599  ThrValProGluHisProHisProHisProHisProHisProHisProHisProHis 1618
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      62      ACCGTGCCGAGCAGCACCCACACCCCTATCGCCCTATGAGCACCTGCTTCGGGGCGTG 121
Qy      1619  SerGlyValAepLeuTyArgSerHisProLeuAlaPheAspProThrSerIlePro 1638
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      122      AGTGGCGTGGACCTGTATCGAGCGACATCCCTCTGGCCTTCGACCCCACTCCATACCC 181

```

```

Qy      1639  ArgGlyIleProLeuAepAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 1658
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      182      CGCGGCATCCCTCTGGACGACGCGCTGCTTACTACCTGCGCCGACACCTGGCCCCCAAC 241
Qy      1659  ProThrTyProHisLeuTyProProTyrieulleArgGlyTyProAspThrAlaAla 1678
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      242      CCCACCTACCGCACCTGTACCCACCTTACCTATCCCGGCTACCCGACACCGCGGGCG 301
Qy      1679  LeuGluAenArgGlnThrIleIleAsnAspTyrieulleThrSerGlnGlnMetHisHisAsn 1698
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      302      CTGGAGAACCGGCAGACCATCATCAATGACTACATCCTCGCAGCAGATGCACCAAC 361
Qy      1699  ThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSer 1718
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      362      GCGGCCACCGCATGCGCCAGCAGCTGATATGCTGAGGGGCTCTCGCCCCCGAGTCC 421
Qy      1719  SerLeuAlaLeuAsnTyAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValPro 1738
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      422      TCGCTGGCACTCAACTACGCTCGGGTCCCGAGGATCATCGACCTGTCCCAAGTGCA 481
Qy      1739  HisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeu 1758
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      482      CACCTGCTGTGCTCGTGCCTCCGACACCCAGGACCCCGCCATCGACCTGTCCCAAGTCC 541
Qy      1759  AlaTyLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSer 1778
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      542      GCCTACCTNCCACCGCGCCCGCCAGCCCTTCAGAGCGCGCCAGCAGCTGCCCATCTTCC 601
Qy      1779  ProGlyGlyProHisLeuThrIleTyProThrThrThrSerSerSerGluArgGluArg 1798
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      602      CCAGGAGGTCCACACACTTGAC-AAAACCAACCAACCAACCAACCAACCAACCAACCA 660
Qy      1799  Asp-ArgAspArgGluArgAspArgGluArgGluArgGluArgGluArgGluArgGluArg 1818
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      661      GACCGGGATCGAGAGCGGACCGGATCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Qy      1818  rThrThrValGluHisAlaProIlePheArgProGlyThrGluGlnSerSerGlySerSe 1838
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      721      CACGACGCTGGAGCAGCAGCCCATCTGGAGACCTGTACAGACANAGCANNCAGCAGCAG 780
Qy      1838  rGlySerSerGlyGlyGlyGlyGlySerSerSerSerSerSerSerSerSerSerSerSer 1857
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      781      CGGAGCAGCGGGGGTGGGGGGGGCGCCCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 823
Qy      1857  aHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnArgProSerVa 1877
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      824      -----TTTCCCGTT-----CCCGGC----- 838
Qy      1877  lLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrVa 1897
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      839      -----TTCTCNCCTGTTTNTCTGCGCGTGTGCGCTCTCTCGT 873
Qy      1897  lLeuArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProAlaAlaTh 1917
      |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      874      TCTC-----TTTGTTCCTCTCCCTTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 924
Qy      1917  rHisCysProLeu 1921
      : |||||
Db      925      TTCTTCGCGCCCTT 937

RESULT 36
BX172348
LOCUS      BX172348
DEFINITION      AGENCOURT_7970482 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6164778
5', mRNA sequence.
ACCESSION      BX172348
VERSION      BX172348.1 GI:22686332
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 898)

```

AUTHORS NIH-MGC <http://mgc.mci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1522 row: g column: 19
High quality sequence stop: 593.
FEATURES Location/Qualifiers
source
1..898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6164778"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 8,296-45 Length: 898
Score: 1305.50 Matches: 265
Percent Similarity: 86.32% Conservative: 0
Best Local Similarity: 86.32% Mismatches: 17
Query Match: 9.88% Indels: 25
Dbs: 13 Gaps: 6

US-09-522-753-5 (1-2517) x BUI172348 (1-898)

QY 1462 GlyAlaSerThrThr-GlySerLysLysHisAspValArgSerLeuIleGlySerProGly 1481
Db 3 GCGCGGTCCACACCTGGGGCTCCAAAGACGACGATGCTCCCTCATCGGAGCCCCGG 62
QY 1481 YArgThrPheProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGluAr 1501
Db 63 CCGGACGTTCCACCGGTGACCCGCTGGATGTGATGCGCGAGCGCCGCGACCTGGAACG 122
QY 1501 gAlaCysTyrGluLysSerLeuLysSerArgProGlyThrAlaSerSerGlyGlySe 1521
Db 123 TGCTTGCTACGAGGAGAGCCTGAAGACCGCGCCAGGACCGCCAGCAGCTCGGGGGCTC 182
QY 1521 rIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerProLe 1541
Db 183 CATTCGCGCGCGCGCCCGCTGATGTGCTGAGTGGGTAAAGCGCGGACAGCCCCCT 242
QY 1541 uThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProValTh 1561
Db 243 GACCTATGAGGACCAACGCGGACCCCTTTGCGCGCCACCTCCACAGAGTTCGCCGTGAC 302
QY 1561 rMetArgGluProThrProArgLysGlnGluGlySerLeuSerSerLysAlaSerG1 1581
Db 303 CACGCGGAGGCCACGCGCGCTGACGAGGGGAGCCTTTGTCGACAGAGGATCCCA 362
QY 1581 nAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrValPr 1601
Db 363 GGACGAAGCTGAGCTGACGCTCGTGAGATCCCAAGTCCCGGACAGACCGCTGCC 422
QY 1601 oGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGlyVa 1621
Db 423 CGAGCACCAACACACACCCCACTCCCTATGAGCACTGCTTCGGGGCGTGAGTGGGT 482
QY 1621 lAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGlyI1 1641

Db 483 GGACCTGTATCGACGCCACATCCCCCTGGCCTTGGACCCCACTCCATACCCCGCGCAT 542
QY 1641 eProLeuAspAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTy 1661
Db 543 CCCTCTGACGCGCGCTGCTACTACTCTGCCCGACACCTGGCCCCCAACCCACCTA 602
QY 1661 rProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAs 1681
Db 603 CCGCACCTGTACCCACCTACCTCATCCGGGGTACCCCGACACGCGGGCGCTGAGAA 662
QY 1681 nArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaTh 1701
Db 663 ACGGACGACCATCATCAATGACTATACCTCGGACGAGATGACCAACACGCGCCAC 722
QY 1701 rAlaMetAla---GlnArgAlaAspMetLeuArgGlyLeuSer---ProArgGluSerSe 1719
Db 723 CGCCCATGGCCCCCAGAGCTGGATATGGCTGAGGGGGCCCTCTCCCCCGCGGAGTCTCTC 782
QY 1719 rLeu-AlaLeuAsnTyrAlaAla-----GlyProA 1729
Db 783 CTGGGCACTCAACTACCTACCCCTCGCGGTCCCGCGAGGCATTTCATCGAACCGGTCCCCAA 842
QY 1729 rGlyIleIleAspLeuSerGlnValProHisLeuProValLeuVal-----Prop 1746
Db 843 GGGGC-----CCCCAC---CCTGCCCTGGTGTCTCGGGGGCCCCC 878
QY 1746 roThrProGlyThrPro 1751
Db 879 CCAACCCCGCGGGCCCC 895

RESULT 37
BM560255
LOCUS BM560255 1045 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT 6563840 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744688
5', mRNA sequence.
ACCESSION BM560255
VERSION BM560255.1 Gi:18804529
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC <http://mgc.mci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12766 row: h column: 01
High quality sequence start: 18
High quality sequence stop: 616.
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744688"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for

FEATURES
source

full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN		US-09-522-753-5 (1-2517) x BM560255 (1-1045)	
Alignment Scores:		Pred. No.: 1.03e-44 Length: 1045	
Score:		1305.00 Matches: 267	
Percent Similarity:		87.95% Conservative: 3	
Best Local Similarity:		86.97% Mismatches: 31	
Query Match:		9.88% Indels: 6	
DB:		12 Gaps: 2	
US-09-522-753-5 (1-2517) x BM560255 (1-1045)			
QY	1146	GlyPro-----ValThrMetGlyLeuProLeuProMetAspProLysLeuAla	1162
DB	6	GGTCCGGAATTCCTTGATCCATCGGGCTGCCCTGCCATGAGCCCAAAAGCTGGCA	65
QY	1163	ProPheSerGlyValLysGlnGluLeuSerProArgGlyGlnAlaGlyProGlu	1182
DB	66	CCCTTCAGCGAGTGAACAGAGCAGCTGTCCACGGGGCCAGGCTGGGCCACCGGAG	125
QY	1183	SerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeuGlySer	1202
DB	126	AGCTGGGGGTGCCACAGCCAGAGCGCTCGGTGCTGAGGGGACAGCTCTGGGTCA	185
QY	1203	ValProGlyGlySerIleThrLysGlyLeuProSerThrArgValProSerAspSerAla	1222
DB	186	GTTCGGGGCGGAAGCATCACAAAGCATTCACAGCACACGGGTGCCCTCGGACAGCGCC	245
QY	1223	IleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThr	1242
DB	246	ATCATACACCGGGCTCCATACCCACGCGCAGCCGCTGACGCTCTGTACAGGGCACC	305
QY	1243	IleThrArgIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSer	1262
DB	306	ATCACAGGATCATCGGGGAGGACAGCCGAGTCTGGACCGCGCGGGAGGACAGC	365
QY	1263	LeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGly	1282
DB	366	CTGCCCAAGGGCCAGCTCATCTACGAAGCAAGAGGGCCAGCTTGTCTATGAGGGT	425
QY	1283	GlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerGlyProProHis	1302
DB	426	GGATGTCTGTACCCAGGCTCTCCAGAGGAGCGCAGACAGCTCAGACCCCCCAT	485
QY	1303	GluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIle	1322
DB	486	GAGAGGGCGGCCCAAGCGCACCTATGACATGATGGAGGGCGCGTGGGAGAGCCATC	545
QY	1323	SerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHisSerPro	1342
DB	546	TCCTCAGCAGCATCGAAGGTCTCATGGCCGTGTCATCCCGCGGAGGACAGACGCC	605
QY	1343	HisHisLeuLysGlnHisHisIleArgGlySerIleThrGlnGlyLeuProArgSer	1362
DB	606	CACCACCTTAAGAGCAGCAGCACCATCCCGGGGTCCATCACAAAGGATCCCTCGGATC	665
QY	1363	TyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeu-LeuLysArgGluGly	1382
DB	666	TACGTGGAGCAGCAGGAGACTACTTGTCTGGGAGGCCAAGCTCCCTAAAGCGGAGGG	725
QY	1382	YThrProProProProSerAspLeuThrGluAlaTyrLysThrGlnAlaLeu	1402
DB	726	NACGNTCCGACCCCATCGCCCTCAGGACCTGACCCGAGGCTTACAAACCGAGCCCC	785
QY	1402	uGlyProLeu---LysLeuLys-ProAlaHisGluGlyLeuValAlaThrValLysGluA	1421
DB	786	TGGGCCCCCTGAACCTGAACCCGCGCCATGAAGGGCCCTGGTGGCCCGACGGTGAAG	845
QY	1421	laGlyArgSerIleHisGluIleProArgGluGlnLeuArgHisThrProGluLeuProL	1441

DB	846	GAGGGCGGGCCCGCTCCCATCATGAGATCCCGCTCCAGGAAACTCGGGACCAACCC	905
QY	1441	euAlaProArgProLeu 1446	
DB	906	CCCAACCTGCCCCCTG 922	
RESULT 38			
BO691710			
LOCUS			
DEFINITION			
AGENCOURT_8046381 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209388			
5', mRNA sequence.			
ACCESSION			
VERSION			
BO691710.1 GI:21817026			
KEYWORDS			
EST.			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1 (bases 1 to 875)			
NIH-MGC http://img.nci.nih.gov/.			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-remail.nih.gov			
Tissue Procurement: ATCC			
cDNA Library Preparation: Rubin Laboratory			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L1CM2368 row: j column: 13			
High quality sequence stop: 551.			
FEATURES			
source			
1..875			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:6209388"			
/tissue_type="ductal carcinoma, cell line"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_110"			
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;			
Site 2: EcoRI; cDNA made by oligo-dT priming.			
Directionally cloned into EcoRI/XhoI sites using the			
following 5' adaptor: GGCAGAG(G). Library constructed by			
Ling Hong in the laboratory of Gerald M. Rubin (University			
of California, Berkeley) using ZAP-cDNA synthesis kit			
(Stratagene) and Superscript II RT (Life Technologies).			
Note: this is a NIH_MGC Library."			
ORIGIN			
Alignment Scores:			
Pred. No.: 1.62e-43 Length: 875			
Score: 1274.00 Matches: 255			
Percent Similarity: 89.90% Conservative: 3			
Best Local Similarity: 88.85% Mismatches: 17			
Query Match: 9.64% Indels: 12			
DB: 13 Gaps: 1			
US-09-522-753-5 (1-2517) x BO691710 (1-875)			
QY	1334	AlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGly	1353
DB	1	GCCATCCCGCGGAGCAGACAGCCCCCACCTCAAGAGAGCAGCACCATCCGCGGG	60
QY	1354	SerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArg	1373
DB	61	TCCATCACACAGGAGATCCCTCGGTCTCTACGTGAGGACAGAGGAGGACTACTCGGTCGG	120
QY	1374	GluAlaLysLeuLeuLysArgGluGlyThrProProProProProProProProProPro	1393
DB	121	GAGGCCAAGCTCTTAAAGCGGGGAGGAGCGCTCCGCGCCCGCCCGCTCAGGGACCTG	180

```
QY 1394 ThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGly 1413
Db 181 ACCGAGGCTACAGACGAGCCCTGGGCCCTTGAAGCTGAAGCCGGCCCATGAGGCG 240
QY 1414 LeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeu 1433
Db 241 CTGGTGGCCACGGTGAAGAGGCGGGCGCTCCATCCATGATGATCCGCGAGAGGTG 300
QY 1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGln 1453
Db 301 CGGCACAGCCCGACGTGCGCCCTGGCCCGCGCCGCTCAAGAGAGGCTCCATCAGCAG 360
QY 1454 GlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
Db 361 GGCACCCCGCTCAAGTACGACACCGCGCGCTCCACCTGGCTCCAAAAGCAGCAGCTA 420
QY 1474 ArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMet 1493
Db 421 CGCTCCCTCATCGGAGCGCCCGCGGAGCTTCCACCCGCTGCACCCGCTGGATGTGATG 480
QY 1494 AlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGly 1513
Db 481 GCCGACGCCCGGCGCACTGGAAGCTGCTGCTACGAGAGAGCTGAAGAGCGCGCAGGG 540
QY 1514 ThrAla-SerSerSerGlySerIleAlaArgGlyAlaProValIleValProGluLe 1533
Db 541 ACCGCCAGNAGCTCGGGGGGCTCCATGGCGCGCGCGCCCGGTCATTGTGCTGAGCT 600
QY 1533 uGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHi 1553
Db 601 GGGTAAGCGCGGCGAGAGCCCTGACCTATGAGGACCAACGCGGCGACCTTTGCGCGCA 660
QY 1553 sLeuProArgGlySerProValThrMetArgGluProThrProArg-LeuGlnGluGlyS 1573
Db 661 CTTCCACAGAGTTCCCGCGTGACCAACGCGGAGCCCGCGCGCTGAGGAGGGCA 720
QY 1573 erLeuSerSer-SerLysAlaSerGlnAsp-ArgLysLeuThr----- 1586
Db 721 GCCTTTCGTCAGGCAAGCATCCCAAGAACCGAAAGCTGACCGTCGAGCCCTCCGTGA 780
QY 1587 -----SerThrProArgGluIleAlaLysSerProHisSerThrValProGluHisHi 1604
Db 781 GAATCGCAAGGTTCCTCCGCGCACAGCAACCGGCGCGCGGAAACCAACCAACCCCA 840
QY 1604 sProHisProfile 1608
Db 841 TTCTTCGCCCTTA 853
```

```
RESULT 39
CF272442
LOCUS CF272442 915 bp mRNA linear EST 13-AUG-2003
DEFINITION AGNCOURT_15178952 NIH_MGC_192 Homo sapiens cDNA clone
IMAGE:30512556 5', mRNA sequence.
VERSION CF272442
KEYWORDS CF272442.1 GI:33628354
SOURCE EST.
ORGANISM Homo sapiens (human)
```

```
REFERENCE 1 (bases 1 to 915)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gethard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaaps-re@mail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Agencourt
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAMS73 row: a column: 13
High quality sequence stop: 627.
Location/Qualifiers
1. 915

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30512556"
/tissue_type="Brain glioblastoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 192"
/note="Vector: pExpressi; Site 1: SmaI; Site 2: NotI; The
library was constructed by reverse transcription of 1 ug
mRNA using the oligo dt primer GCGGCCGCC(T)20 and an
RNaseH + MMLV reverse transcriptase. Second strand
synthesis was carried out by standard methods. The cDNA
was size selected by agarose gel for > 1.2 kb, digested
with Not I and directionally cloned into the vector
Express-1 at the SmaI/NotI sites. DNA from the primary
library was used for in vitro transcription from the T7
promoter to produce biotinylated RNA transcripts. These
biotinylated transcripts, along with blocking oligos to
the poly-A, multiple cloning site and primer regions, were
hybridized with single stranded circles produced by
phageid production from the primary library to a Cot
value of 10-20. Streptavidin/phenol extraction was utilized
to remove DNA:RNA hybrids leaving un-hybridized single
stranded circles which were repaired by primer extension
and transformed back into E. coli resulting in the
normalized library. Average insert size 2.0 kb. 3'
linker/adaptor sequence GCGGCCGCC(T)20. This library was
constructed by Agencourt Bioscience."

ORIGIN

Alignment Scores:	Pred. No.:	Length:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
	5,18-43	Matches:	915	77.40%	74.86%	9.55%	
	1262.50	Conservative:	265				
		Mismatches:	9				
		Indels:	17				
		Gaps:	4				

US-09-522-753-5 (1-2517) x CF272442 (1-915)

QY	2149	ArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGly	2168
Db	3	CGGCACCAACCCACAGCAGCTCAGCGCACCTGCGCGCCCTCTACTCTCTCTCTGGG	62
QY	2169	AlaSerCysProValLeuAspLeuArgProProSerAspLeuTyrLeuProProPro	2188
Db	63	GCCAGCTGCCCGCTCTGGACCTCGCGCGCCACCCAGTGACCTTACCTCCCGCCCG	122
QY	2189	AspHisGlyAlaProAlaArgGlySerProHisSerGluGlyLysArgSerProGlu	2208
Db	123	GACCATGTGTCGCCCGCGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	182
QY	2209	ProAsnLysThrSerValLeuGlyGlyGluAspGlyIleGluProValSerProPro	2228
Db	183	CCAAACAAGACGTCTGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	242
QY	2229	GluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAsp	2248
Db	243	GAGGCGCATGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	302
QY	2249	GlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnPro	2268
Db	303	GGGGAACAGACGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	362
QY	2269	ProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysGln	2288

FEATURES		source		High quality sequence start: 2		High quality sequence stop: 549.	
Location/Qualifiers		1..1075		/organism="Homo sapiens"		/mol_type="mRNA"	
				/db_xref="taxon:9606"		/clone="IMAGE:5501184"	
				/tissue_type="retinoblastoma"		/lab_host="DH10B (phage-resistant)"	
				/clone_lib="NIH_MGC_67"		/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Library constructed by Life Technologies."	
ORIGIN		Alignment Scores:		Pred. No.:		Length:	
				Score:		Matches:	
				Percent Similarity:		Conservative:	
				Best Local Similarity:		Mismatches:	
				Query Match:		Indels:	
				DB:		Gaps:	
				US-09-522-753-5 (1-2517) x BM461469 (1-1075)			
Db	363	CCAGCCTTCTTTCAGACAGCTGACCGAGACGAACTCCGCCATGGTCAAGTCAAGAGACAA	422				
Qy	2289	Glut1eLeuLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGln	2308				
Db	423	GAGATCAACAAGAGCTGAACACCCACACCGGATGAGCTGATACATATACAGCCAG	482				
Qy	2309	ProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArg	2328				
Db	483	CCTGGGCGGAGATCTTCAATATGCTCCGCCATCACCGGAACAGGCTTATGACCTAGA	542				
Qy	2329	SerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLys	2348				
Db	543	AGCCAGGCGGTGCAGGAACATGCCAGCACCAATGGGGCTGGAGGCATATATAGAAAG	602				
Qy	2349	AlaLeuMetGlyLysTyrAspGluTrpGluSerProLeuSerAlaAsnAlaPhe	2368				
Db	603	GCACATAG-----	611				
Qy	2369	AsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGly	2388				
Db	611	-----	611				
Qy	2389	ArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArg	2408				
Db	612	-----GGTGGCGCGGGAAGGCAAGGTCTCTGGCAGA	644				
Qy	2409	ProSerSerArgLysAlaLysSerProAlaProGly-LeuAlaSerGlyAspArgPro-P	2428				
Db	645	CCACAGCGCGAAAGCAACCAAGTCCCGCGCCCGCTGGCATCTGGGACCGGCCACC	704				
Qy	2428	roSerValSer-SerValHisSer-GluGlyAspCysAsnArg-----ArgThrProLe	2445				
Db	705	CCTCTGTCTCTCCATGTCAGTCCGGAGGGAGACTGCAACCGCCCGGACCGCGCTCAAC	764				
Qy	2445	uThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAs	2465				
Db	765	CAACCGCGGTGTGGGGAAGAGGCGCCCTTCGTCGGCAGGTTTCCACCGCCATTTCC	824				
Qy	2465	nPro-----LeuIleMetArgLeuGln--AlaGlyValMet-----	2476				
Db	825	CCCTTCAACACCCCGCTGATCATGCGCGCTGACGGGCGGGGTGTCAATGGGNTT	884				
Qy	2477	---AlaSerProProProProGlyLeuPro 2485					
Db	885	TTCCCCACACCCCGCCCGCGGGGCCCT 914					
RESULT 40		BM461469		1075 bp		mRNA	
LOCUS		AGENCOURT_6420985 NIH_MGC_67		Homo sapiens		cDNA clone IMAGE:5501184	
DEFINITION		5', mRNA sequence.					
ACCESSION		BM461469					
VERSION		BM461469.1		GI:18510509			
KEYWORDS		EST.					
SOURCE		Homo sapiens (human)					
ORGANISM		Homo sapiens					
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE		NIH-MGC http://mgi.nci.nih.gov/.					
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT		Unpublished (1999)					
		Contact: Robert Strausberg, Ph.D.					
		Email: egapbs-x@mail.nih.gov					
		Tissue Procurement: ATCC					
		cDNA Library Preparation: Life Technologies, Inc.					
		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
		DNA Sequencing by: Agencourt Bioscience Corporation					
		Clone distribution: MGC clone distribution information can be					
		found through the I.M.A.G.E. Consortium/LLNL at:					
		http://image.llnl.gov					
		Plate: LLAM12137 row: f column: 01					

FEATURES	Location/Qualifiers																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
----------	---------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

QY 448 ProLysAspPheGlyLeuIleAlaSerPheLeuGluArgLysThrValAlaGluCysVal 467
 DB 124 CCCAAGAACTTTGGCCTGATCGCATCTTCTGGAGAGAAACAGTGGCTGAGTGGCTC 183
 QY 468 LeuTyrTyrLeuTyrLysLysAsnGluAsnTyrLysSerLeuValArgSerTyr 487
 DB 184 CTCTATTACTTACTGACTAAGAGAAATAGAACTAATAGAGCTTGGTGAGACGGAGCTAT 243
 QY 488 ArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 507
 DB 244 CGGCC-CGCGCAAGAGCCAGCAGCAACACAGCAGCAGCAGCAGCAGCAGCAGCAG 302
 QY 508 GlnGlnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGluLys 527
 DB 303 CAGCAGCAGCCATGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362
 QY 528 GluAlaGluLysGluGluGluLysProGluValGluAsnAspLysGluLysLeuLys 547
 DB 363 GAGCGGAG 422
 QY 548 GluLysThrAspThrSerGlyGluAspAsnAspGluLysGluAlaValAlaSerLys 567
 DB 423 GAGAGACAGACACACCTCAGCGGAGAGACAGCAGCAGCAGCAGCAGCAGCAGCAG 482
 QY 568 GlyArgLysThrAlaAsnSer-GlnGlyArgArgLysGlyArgIleThrArgSerMetAl 587
 DB 483 GCGCGCAAACTGCCAACAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 542
 QY 587 AsnGluAlaAsnSerGlu-GluAlaIleThrProGlnGlnSerAlaGluLeuAlaSerM 607
 DB 543 TAATGAGGCCAACAGCAGAGAGGCCATCACCCTCCAGCAGCAGCAGCAGCAGCAGCAG 602
 QY 607 etGluLeuAsnGluSerSerArgTTPThrGluGluMetGluThrAlaLysLysGlyL 627
 DB 603 TGGACTGATGATGAGTTCCTCTGGACAGAAAGAAATGAAACCCGCCAGAAAGTTC 662
 QY 627 euLeuGluHisGlyArg-AsnTrpSerAlaIleAlaArg-MetValGlySerLysThrVa 646
 DB 663 TCCTGGACACAGCCGCCAACTGGCGGCCATCCCGCGGATGGGGGCTCCCAACTGG 722
 QY 646 Lser-GlnCys-LysAsnPheTyrPheAsnTyrLysArgGluAsnLeuAspGluLe 665
 DB 723 GTCCCGCAGTGTAGAACTTCTACTTCACTACAG-AAGAGCAGCAGCAGCAGCAGCAG 781
 QY 666 LeuGlnGlnHisLysLysMetGluLysGluArgAsnAlaArgArgLysLysLysLys 685
 DB 782 TTGACGCACACAG-CTGAAAAGGAAAGAGAGAGAACCCCGGAGAAACCAAGAAA 840
 QY 686 AlaProAlaAlaAlaSerGluGluAlaAlaPheProProValValGluAspGluGlu-Me 705
 DB 841 GCGCCCGGGGGGCCAACCAAGAGAGCTGCATTCCCGCGGGGGAAGATAGGGAAG 900
 QY 705 tGluAla-SerGlyValSerGlyAsnGluGluMetValGluGluAlaGluAlaLeuH 725
 DB 901 GGAGCCCCCGGGCGCGCGAGATGAGACGAAAAATGTT----- 941
 QY 725 leAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAlaThrValAsnAsn 745
 DB 942 -----GGAACAGAGTCT-----TCATACCCCTGT---GTCCAAAATG 975
 QY 745 er-SerAspThrGluSerIleProSerProHisThrGluAlaAlaLysAspThrGlyGln 764
 DB 976 CTCTCAACCGGGGACATTCCTCTCTCCACGGGGGGCGCCCGGAGGCGGCGGCTC 1035
 QY 765 AsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProProGly 781
 DB 1036 AATGGT-----GGGACAACTC-CCACCCCGCGGG 1064

RESULT 44

CB723733

LOCUS

DEFINITION

797 bp mRNA linear EST 09-JUL-2003
 UI-M-FY0-cez-b-23-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone

IMAGE:6844056 5', mRNA sequence.

CB723733

CB723733.1 GI:29780875

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 797)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9pbbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/mousefl.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5

Location/Qualifiers

1..797

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6844056"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_FY0"

note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is AGCAGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1,14e-42 Length: 797

Score: 1252.50 Matches: 239

Percent Similarity: 93.23% Conservative: 9

Best Local Similarity: 89.85% Mismatches: 16

Query Match: 9.48% Indels: 2

DB: 14 Gaps: 1

US-09-522-753-5 (1-2517) x CB723733 (1-797)

QY 1480 ProGlyArgThrPheProValHisProLeuAspValMetAlaAspAlaArgAlaLeu 1499
 DB 2 CCGCGCGCGCTTCCCTCCCTGCCCTGACATANTGGTCAGCCCGGCGCATG 61

QY 1500 GluArgAlaCysTyrGluSerLeuLysSerArgProGlyThrAlaSerSerSergly 1519
 DB 62 GAGCGTGCCTCTATGAGAGAGTCTGAAGAGCGGTCCAGGACCCAGCAGTGGTGCAGGG 121

QY 1520 GlySerIleAlaArgGlyAlaProGluLeuValProGluLeuGlyLysProArgGlnSer 1539
 DB 122 GGCTCCATCATACAGTGGGGCTCCAGTCGTGCTGAACCTGGGCAAGCCAGCCGGAAGC 181

```
Qy 1540 ProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerPro 1559
Db 182 CCACTGACTACGAGAGACCCAGCGGCGACCTTACAGTACCTGCGCAGGTGGCTCCCT 241
Qy 1560 ValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAla 1579
Db 242 GTGACCAAGAGGAGCCACGCCAGCTTTCAGGAAGCAGCCTTCTATCCAGCAAGCG 301
Qy 1580 SerGlnAspArgLysLeuThrSerThrProArgGluLeuAlaLysSerProHisThr 1599
Db 302 TCCAGGAGCCGGAAGCTGACATCTACACCCGGGAGATCGCAATGCCACACAGCACT 361
Qy 1600 ValProGluHisHisProHisProHisProHisProHisProHisProHisProHis 1619
Db 362 GTGCCGAGACCCACCTTACCCCATCTCCCTATGAGCATTTGCTCCGGGGGTGACT 421
Qy 1620 GlyValAspLeuTyrArgSerHisLeuProLeuAlaPheAspProThrSerIleProArg 1639
Db 422 GGTGTGGACCTGTACCGTGGTCAATCCATTTGACCCACCTCCATACCCCGA 481
Qy 1640 GlyIleProLeuAap--AlaAlaAla--TyrTyrLeuProArgHisLeuAlaProAs 1658
Db 482 GGGATCCCTCTGGAGAGCAGCGCAGCTTACTTACCTGCCCGGCACTTGGGCCCGAG 541
Qy 1658 nProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAl 1678
Db 542 CCCCACCTTACCAACCTGTACCCACCTTACTTATCCCGGCTACCTGACACGGCGGC 601
Qy 1678 aLeuGluAsnArgGlnThrIleAlaAsnAspTyrIleThrSerGlnGlnMetHisHisAs 1698
Db 602 CCGTGAGAAACGCCAGACCATCATCAATGACTATACATCACTCGGAGAGATGCCACAA 661
Qy 1698 nThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSe 1718
Db 662 CGCTGCCCTCGCCATGGCCGCGGCTGATGTCATGCTGAGGGGTCTGTACCGGAGAGTC 721
Qy 1718 rSerLeuAlaLeuAenTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValPr 1738
Db 722 CTCGTGGNCTCAATTATGCGCTGGCCGAGGAGCATTCGACCTGTCCCAAGTGC 781
Qy 1738 oHisLeuProValLeu 1743
Db 782 ACACCTGCCGCTGTG 797

RESULT 45
BG252257
LOCUS
DEFINITION
602365136F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473272 5',
mRNA sequence.
ACCESSION
BG252257
VERSION
BG252257.1 GI:12762073
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteleostomi; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1087)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0295 row: h column: 09
High quality sequence stop: 641.
Location/Qualifiers
1..1087
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4473272"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores: 2,03e-42 Length: 1087
Pred. No.: 1250.00 Matches: 276
Score: 75.65% Conservative: 13
Percent Similarity: 72.25% Mismatches: 61
Best Local Similarity: 9.46% Indels: 33
Query Match: 12 Gaps: 4
DB: 12

US-09-522-753-5 (1-2517) x BG252257 (1-1087)

Qy 1091 ThrAlaArgProValLeuProArgProThrIleSerAsnProProLeuIleSer 1110
Db 2 ACTGCCGCGCGCTCTGCTGCGCGCCACCCACCATCTCCAAACCGCGCTCCCTCATCTCC 61
Qy 1111 SerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSer 1130
Db 62 TCTGCCAAGCAGCCCGAGCGTCTCGAGAGGCAATAGTGCCATCTCCCAAGGAATGTG 121
Qy 1131 ValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMet 1150
Db 122 GTCAGCTCCACGCTCCGCTACTCAGACATGCCAGCCCGCGTGGCCCTGTCCACATG 181
Qy 1151 GlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGlu 1170
Db 182 GGGCTGCCCTTGGCCATGGACCCCAAAAGCTGACACCTTCAGCGGAGTGAAGCAGAG 241
Qy 1171 GlnLeuSerProArgGlyGlnAlaGlyProGlySerLeuGlyValProThrAlaGln 1190
Db 242 CAGCTGTCCCGACGGGCGCCAGGCTGGGCGCCAGAGAGCTGGGGGTGCCACAGCCAG 301
Qy 1191 GluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLys 1210
Db 302 GAGCGCTCCGCTGAGAGGAGCAGCTCTGGGCTCAGTTCGGGCGGAGCATCACCAA 361
Qy 1211 GlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThr 1230
Db 362 GGCATTCCAGCACACGGGTGCCCTCGGACAGCGCCATCATATACCGCGGCTCCATCACC 421
Qy 1231 HisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAsp 1250
Db 422 CACGCGCAGCGCAGCTGAGCTCTCTAAGGGACCATCATCAGAGATCATCGGCGAGAGC 481
Qy 1251 SerProSerArgLeuAspArgGlyArgGlu-AspSerLeuProLysGlyHisValIleTy 1270
Db 482 AGCCCGAGTGCCTTGGACCGCGCGGACGAGCAGCTGCCCAAGGCCACGATCATCTA 541
Qy 1270 rGluGlyLysGlyHisValLeuSerTyrGlu-GlyGlyMetSerValThrGlnCys 1290
Db 542 CGAAGGCAAGAGGGCCACCGTCTTGTCTATACGGGTGGCATGTCTGTGACCCAGTGCT 601
Qy 1290 erLysGluAspGlyArg-SerSerSerGlyProHisGluThrAlaAlaProLysArg 1309
Db 602 CCAAGGAGGAGCGGAGAAAGCAGCTCAGGACCCCGCCATGAGCCGCGCCCCCAAGCGC 661
Qy 1310 ThrTyrAspMetMet-GluGlyArgValGly---ArgAlaIleSerSerAla--SerIle 1327
Db 662 ACCTATGACATGATCGAGGGCCGCGTGGGCGACGAGCCATCTCTCAAGCCCAAGCATC 721
Qy 1328 GluGlyLeuMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGlu 1347
```

```

Db      722  GAACTCTCATGGCGCGCC-ATCCCGCGGAGCAGACAGCCGCCACCACTCAACAGAC 780
Qy      1348  GlnHisHisIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGln 1367
Db      781  CACCAACCAATCCGGGGTCCATCACCCAGAGT---CCCGGTCTTCGTGTGAGACCGGA 837
Qy      1368  GluAspTyrLeuArgGluAlaLysLeuLeuLysArgGluGlyThrProProPro 1387
Db      838  AGA---GCAACTGGATGGGAGCAGATCTTAAGCGCGGACACACCCCTCCCGCCGCCA 894
Qy      1388  ProProSerArgAspLeuThrGluAlaTyrIysThrGlnAlaLeuGlyProLeuLysLeu 1407
Db      895  CAGGCTCTCCATGACATATAACACCGCGATCCCAATAAATACTACCCCACTACACTC 954
Qy      1408  LysProAlaHisGluGlyLeuValAlaThrValIysGluAlaGlyArgSerIleHisGlu 1427
Db      955  TTA----- 957

Qy      1428  IleProArgGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLys 1447
Db      958  -----CAAAAAGGCACACTAATTAATCACTCACTACTCACACGCTCTCTCCG 1005

Qy      1448  GluGlySerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThr 1465
Db      1006  CAACATCTCCGCTCACACACATTAACCTCTCATCTATCCACACCTCAACGCTACC 1059

RESULT 46
BG831424
LOCUS   602766347F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4908491 5',
DEFINITION mRNA sequence.
ACCESSION BG831424
VERSION   BG831424.1 GI:14179011
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 958)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1814 row: n column: 12
High quality sequence stop: 914.
Location/Qualifiers
1. 958
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4908491"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 42"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.44e-42 Length: 958
Score: 1243.00 Matches: 275
Percent Similarity: 88.47% Conservatives: 9
Best Local Similarity: 85.67% Mismatches: 28
Query Match: 9.41% Indels: 15
DB: 12 Gaps: 5

US-09-522-753-5 (1-2517) x BG831424 (1-958)

Qy      1292  GluAspGlyArgSerSerSerGly-ProProHisGluThrAlaAlaProLysArgThrTy 1311
Db      2      GAGGACGCGCAGACAGCTCAGGACCCCGCATGAGACGCGCGCCCAAGCGCACCTA 61
Qy      1311  rAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMe 1331
Db      62  TGACATGATGAGGCGCGGTGGGAGAGCCATCTCTCAGCCAGCATCGATGGTCTCAT 121
Qy      1331  tGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisIle 1351
Db      122  GGGCGGCGC-ATCCCGCGGAGCGACAGACAGCCCGCCACCTCAAGAGCAGCACCAT 180
Qy      1351  eArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLe 1371
Db      181  CCGCGGGTCCATCACACAGGGATCCCTCGGTCTCTACGTGGAGGCACAGGAGTACCT 240
Qy      1371  uArgArgGluAlaLysLeuLysArgGluGlyThrProProProProProProProSer 1391
Db      241  CGGTGGGAGCGCAAGCTCTTAAGCGGAGGAGGACGCGCTCCGCGCCCGCCCTCAGC 300
Qy      1391  gAspLeuThrGluAlaTyrIysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHi 1411
Db      301  GGACTCGACCGAGCGCTACAGACGCGGCGCTTGGGCGCCCTGAAGCTGAAGCGGCCA 360
Qy      1411  sGluGlyLeu-ValAlaThrValIysGluAlaGlyArgSerIleHisGluIleProArg 1431
Db      361  TGAGGGCTGATGATGCCACCGTGAAGGAGGCGCGCGCTCCATCATGAGATCCGCGCG 420
Qy      1431  luGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySer 1451
Db      421  AGGAGCTGCGGCACACGCGCGAGCTGCGCGCGCGCGCGCGCTCAAGGAGGCTCCA 480
Qy      1451  leThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysH 1471
Db      481  TCACGCGAGGCGACCCCGCTCAAGTACGACACCGCGCGCGCTCCACCACTGGCTCCA 540
Qy      1471  isAspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeu 1491
Db      541  ACAGCTACGCTCCCTCATCGCAGCGCGCGAGCGGCGTCCACCGCGCGCGCTGG 600
Qy      1491  spValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSer 1511
Db      601  ATGTGATGGCGGACGCGCGCGGACCTGGAACGTGCTGTCTAGGAGAGCGCGTGAAGAG 660
Qy      1511  rgProGlyThrAlaSerSerGlyGlySerIleAlaArgGlyAlaProValIleValP 1531
Db      661  CGGCGAGGACCGCGCAGCAGCTCGGGGGCTCCATTGGCGCGCGCGCGCTCATTTGTC 720
Qy      1531  roGlu-LeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPhe 1550
Db      721  CTGAGCTTGGGGAAGCGGTGAGAGCCCT---AACTATGAGGACCAACCGGGGACCTTT 777
Qy      1551  AlaglyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGln 1570
Db      778  GC-GGACACTC-CCACGAGGGTCCCGGTG-ACCAGCGGGAGGCCACACG---CGGCTGCA 831
Qy      1571  GluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArg 1590
Db      832  GAAGCGAGCCT-TGCTCCAGCAAGGA-TCCCGA---ACGAGGTGACGCTGACGCTGCT 886
Qy      1591  GluIleAlaLysSerProHisSerThrValProGluHisHisProHisProHisSer 1609

```

```
Db      887 GAGATGCC---AGTCCGACAGA-----CCGGCCGAGACACCAAGACCATTTGGCT 934

RESULT 47
BG974253
LOCUS
DEFINITION 602843992F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4979758 5',
            mRNA sequence.
ACCESSION BG974253
VERSION   BG974253.1 GI:14361890
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 810)
           NIH-MGC http://mgi.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
           Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM10978 row: g column: 23
           High quality sequence stop: 787.
           Location/Qualifiers
             1..810
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="NMRI"
               /db_xref="taxon:10090"
               /clone="IMAGE:4979758"
               /tissue_type="tumor, gross tissue"
               /dev_stage="5 months"
               /lab_host="DH10B"
               /clone_lib="NCI CGAP Mam4"
               /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
               Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
               Library constructed by Life Technologies. Investigators
               providing samples: Lothar Hennighausen/Priscilla Furth,
               NIH Reference for transgenic model: Li et al., Cell Growth
               and Differentiation 7, 3-11 (1996)."
```

ORIGIN

```
Alignment Scores:
Pred. No.:      4,19e-42      Length:      810
Score:          1239.00      Matches:     244
Percent Similarity: 92.4%      Conservative: 1
Best Local Similarity: 92.08%  Mismatches: 17
Query Match:     9.38%      Indels:      3
DB:              12         Gaps:         2

US-09-522-753-5 (1-2517) x BG974253 (1-810)
```

```
Qy      2211 LysThrSerValLeuGlyGlyGluAspGlyIleGluProValSerProGluGly 2230
Db      2 AAAACATCGTCTCGGCAGCAGCAGGATGATGACCTGTGTCCTCCACGAGGGC 61
Qy      2231 MetThrGluProGlyHisSerArgSerAlaValTyProLeuLeuTyArgAspGlyGlu 2250
Db      62 ATGACTGAGCCAGGACATCGGAGCACTGGGTACCCACTGCTGTATCGAGACGGGAA 121
Qy      2251 GlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProAla 2270
Db      122 CAGGGCGAGCCCC---AGGATGGGCTCTAAGTCTCAGGCAACACGAGCGCGCAGCC 178
Qy      2271 PhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysGlnLulle 2290
```

```
Db      179 TTCTTCAGTAAGCTGACTGAGAGCAACTCCGCCATGTTGAGTCGAGAAGCAGGAGATC 238
Qy      2291 AsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyAsnIleSerGlnProGly 2310
Db      239 AACAAAGAACTCAACCCACACACGAGCAGGAGCAAGATACAAATATTCGCAGCCTGGG 298
Qy      2311 ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyArgSerGln 2330
Db      299 ACGGAATCTTCAACATGCCCGCCATCACTGAGGAGCGCTTATGACCTGTAGAACCCAG 358
Qy      2331 AlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeu 2350
Db      359 GCGGTGCAAGAAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAGAAGGCACCT 418
Qy      2351 MetGlyLysTyArgProGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnPro 2370
Db      419 ATGGGTAAATATGATCAGTGGGAGAGAGCCCCCGCGCTGGCGCCAAATGCTTTTAAACCT 478
Qy      2371 LeuAsnAlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyArg 2389
Db      479 CTGAATGCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATACCACTGCTGACGCGCG 538
Qy      2390 SerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArgPro 2409
Db      539 AGTGACCCAGCACTCACTCGCCAGGTGGAGGTGGGAAAGCCAAAGTCTCTGCGACACCT 598
Qy      2410 SerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSer 2429
Db      599 AGCAGCGCAAGAACCAAGTCCGACGACGAGGCTAGGCTCCGGAGACCGACCCCTTCT 658
Qy      2430 ValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgVal 2449
Db      659 GTCTCTCTAGTACACTCAGAGGGGGACTGCAATCGCGCAACCACTTCAACAACCGGTG 718
Qy      2469 tArgLeuGlnAla 2473
Db      779 GAGGCTACAGGCC 791

RESULT 48
BG944732
LOCUS
DEFINITION AGENCOURT_10036107 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481219
           5', mRNA sequence.
ACCESSION BG944732
VERSION   BG944732.1 GI:22360210
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 962)
           NIH-MGC http://mgi.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: DCTD/DTF
           cDNA Library Preparation: Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LICM2662 row: d column: 20
           High quality sequence start: 139
           High quality sequence stop: 597.
           Location/Qualifiers
             1..962
               /organism="Homo sapiens"
```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481219"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/organism="Homo sapiens"
/site="pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 7,42e-42 Length: 962
Score: 1235.00 Matches: 258
Percent Similarity: 81.45% Conservative: 1
Best Local Similarity: 81.13% Mismatches: 6
Query Match: 9.35% Indels: 53
DB: 13 Gaps: 1

US-09-522-753-5 (1-2517) x BQ944732 (1-962)

QY 2207 ProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyLeuGluProValSer 2226
DB 1 CCAGAGCCCAACACGCTGCTTGGTGGTGGTGGAGCGGTATGACCTGTGTCC 60

QY 2227 ProProGluGlyMetThrGluProGlyHisSerArgSerAlaValProLeuLeuTyr 2246
DB 61 CCACCGGAGGCGATCAGCGAGCCACCGGCATCCCGAGGTGCTGTGATCCCGCTGTGTAC 120

QY 2247 ArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSer 2266
DB 121 CGGGATGGGGAACACACGAGCGCCCGCAGGATGGGCTCCAACTCTCCAGGCAACACGAGC 180

QY 2267 GlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLys 2286
DB 181 CAGCGCCGAGCTCTTCAGCAGCTGACCGAGACCACTCCGCGCATGGTCAAGTCCAG 240

QY 2287 LysGlnGluLeuAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIle 2306
DB 241 AAGCAAGAGATCAACAAGAGCTGAACACCCACCAACCGGAATGAGCCTGAATACATATC 300

QY 2307 SerGlnProGlyThrGluLeuPheAsnMetProAlaIleThrGlyThrGlyLeuMetThr 2326
DB 301 AGCCAGCTGGGACGAGATCTTCAATATATGCCCGCATCACCGGAACAGGCTTATGACC 360

QY 2327 TyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIle 2346
DB 361 TATAGACCGAGCGGTGCAGGACATCCAGCACCACTGGGCTGGAGGCCATATTT 420

QY 2347 ArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsn 2366
DB 421 AGAAGGACACTCATG----- 435

QY 2367 AlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaIle 2386
DB 435 ----- 435

QY 2387 AspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyAlaLysValSer 2406
DB 436 -----GGTGGCGCGGGAAGGCCAAGGCTCTCT 462

QY 2407 GlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArg 2426
DB 463 GGCAGACCCAGCAGCCGAAACCAAGTCCCGCGCGCGGCTGGCATCTGGGGACCGG 522

QY 2427 ProProSerValSerSerValHisSerGluGlyAspCysAsnArgThrProLeuThr 2446
DB 523 CCACCTCTGCTCTCAGTGCACTCGGAGGAGACTGCAACCGCGGAGCGCGCTCACC 582

```

```

QY 2447 AsnArgValTrpGluAspArgProSerSerAlaGlySerThr-ProPheProTyrAsnPr 2466
DB 583 AACCGCGTGTGGGAGACAGGCCCTCGTCGCGAGGTTCACGCCCATTCCTCCACACCC 642

QY 2466 OleuLeuMetArgLeuGlnAla-GlyValMetAlaSerPro-ProProProGly-LeuPr 2485
DB 643 CCGTATCATGCGGTGTCAGGGGGGGTGTCAATGGCTTCCCGCCACCCCGGCGCTCC 702

QY 2485 oAlaGlySerGlyProLeuAlaGlyProHisAlaTrpAspGlu-GluProLysProL 2505
DB 703 CGGGGACAGCGGCCCTCGCTGGCCCCCACACAGCTGGGAGGAGGAGGCCGCCAC 762

QY 2505 euLeu-CysSerGlnTyrGluThrLeu-SerAspSerGlu 2517
DB 763 TGCTCTCTGCTCGCAATACGAGACCTCTCTCCGACAGCGAG 802

```

RESULT 49

```

BM811122
LOCUS BM811122 1113 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT 6599082 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5465831
5', mRNA sequence.
ACCESSION BM811122
VERSION BM811122.1 GI:19127945
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1113)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1965 row: d column: 24
High quality sequence stop: 523.

```

FEATURES

```

Location/Qualifiers
1..1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5465831"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.22e-41 Length: 1113
Score: 1231.50 Matches: 259
Percent Similarity: 77.12% Conservative: 14
Best Local Similarity: 73.16% Mismatches: 43
Query Match: 9.32% Indels: 38
DB: 12 Gaps: 7

US-09-522-753-5 (1-2517) x BM811122 (1-1113)

```



```
Qy 1457 LeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeu 1476
Db 1 CTCAAGTACGACACCGGGCGTCCACCACTGGCTCCAAAAGACGACGAGTACGCTCCCTC 60

Qy 1477 IleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAla 1496
Db 61 ATCGGACGCGCGCGGACGCTTCCACCGCTGACCGCTGGATGTGATGGCGGACGCC 120

Qy 1497 ArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSer 1516
Db 121 CGGGACCTGGAACTGGCTCTGACGAGAGAGCTTGAAGACCGCGGACCGGACCGCAGC 180

Qy 1517 SerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysPro 1536
Db 181 AGCTCGGGGGCTCCATTGCGGGCGGCGCCCGGCTCATTTGGCTGAGCTGGGCAAGCG 240

Qy 1537 ArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArg 1556
Db 241 CGGCAGAGCGCGCTAACTATGAGGACCCACGCGGCGACCTTTGGCGGCGACCTCCACGA 300

Qy 1557 GlySerProValThrMetArgGluProThrProArgLeuGlnGlySerLeuSerSer 1576
Db 301 GGTTCGCGCGTGACACCGCGGAGCCACGCGCGCGCTGCGAGGAGGCGAGCTTCGTCC 360

Qy 1577 SerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerPro 1596
Db 361 AGCAAGGATCCAGGACCGAAAGCTGACGTCGACGCGCTGTGAGATCGCCAAAGTCCCG 420

Qy 1597 HisSerThrValProGluHisProHisProLysSerProTyrGluHisLeuLeuArg 1616
Db 421 CACAGCAGCGTGGCGGACGACCCACGACCGGCGACCTTCGCGCTATGAGCAGCTGCTTCG 480

Qy 1617 GlyValSerGlyValAspLeuTyrArgSerHisIleProLeu-AlaPheAspProThrSe 1636
Db 481 GCGGTGAGTGGGTGAGCTGATCGCAGGACACATCCCGCTGGCGCTTCGACCCACCTC 540

Qy 1636 rIleProArgGlyIleProLeuAspAlaAlaAlaTyrTyrLeuProArgHisLeuAl 1656
Db 541 CATACCGCGGCGCATCCCTCTGGAGCGAGCGCTGCTTACTACCTGCGCGGCGGCGGCG 600

Qy 1656 aProAsnProThrTyrProHisLeuTyrProProTyrLeuLeuArgGlyTyrProAspTh 1676
Db 601 CCCCAACCCCACTTACCGCGACCTGTACCCACCTTACCTCTCGCGGCGCTACCCGACAC 660

Qy 1676 rAla-AlaLeuGluAsnArgGlnThrIleAsnAspTyrIle-ThrSerGlnGlnMet 1695
Db 661 GCGCGCGCTGGAGAACCCCGAAACCTCTCTTGTACTTCTTACCTCGACGCGGATG 720

Qy 1696 HisHisAsnThrAla-----ThrAlaMetAlaGlnArgAlaAsp 1708
Db 721 CACCACAAAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777

Qy 1709 MetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyPro 1728
Db 778 CTTCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837

Qy 1729 ArgGlyIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrPro 1748
Db 838 CGAGGA-----AGGCGCCACAC-----CAACTTTGGGCGACACCC 873

Qy 1749 GlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPhe 1768
Db 874 GGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 924

Qy 1769 -----SerSerArgHisSerSerSer 1775
Db 925 CAATGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 984

Qy 1776 ProLeu-----SerProGlyGlyProThrHis 1784
Db 985 CGATTGATTTACTGATGATGGTGGCGCTTACACAC 1020
```

RESULT 50

```
BU191146
LOCUS DEFINITION BU191146 861 bp mRNA linear EST 04-SEP-2002
5' mRNA sequence.
AGENCOURT 7975031 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6081811
BU191146
ACCESSION BU191146
VERSION BU191146.1 GI:22705130
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 861)
REFERENCE NIH-WGAC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2309 row: f column: 20
High quality sequence stop: 550.
FEATURES
Location/Qualifiers
1..861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6081811"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 1.34e-41 Length: 861
Score: 1227.50 Matches: 245
Percent Similarity: 88.61% Conservative: 4
Best Local Similarity: 87.19% Mismatches: 27
Query Match: 9.29% Indels: 6
DB: Gaps: 2

US-09-522-753-5 (1-2517) x BU191146 (1-861)
Qy 1334 AlaIleProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGly 1353
Db 1 GCCATCCCGCGGAGCGACACAGCCCCCACCTCAAGAGCAGCAGCACCATCCGCGGG 60

Qy 1354 SerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArg 1373
Db 61 TCCATCACACAAGGGATCCCTCGGTCTCTACGTGGAGGCACAGGAGGACTACCTCGTCGG 120

Qy 1374 GluAlaLysLeuLeuLysArgGluGlyThrProProProProProProProProProPro 1393
Db 121 GAGCGCCAGCTCTTAAAGCGGAGGCGACGCTCCGCCCGCCCGCCCGCTCAGCGGACCTG 180

Qy 1394 ThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGly 1413
Db 181 ACCGAGGCGCTACAAGCGCAGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

Qy 1414 LeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeu 1433
```

```

Db      241 CTGTGGCCACGGTGAAGAGCGCGCGCTCCATCCATGATGCCCGCGGAGGAGCTG 300
Qy      1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGln 1453
Db      301 CGGCACACGCCGAGCTGCCCTGGCCCGCGCGCTCAAGGAGGCTCCATCAGCGAG 360
Qy      1454 GlyThrProLeuLysThrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
Db      361 GGCACCCCGCTCAAGTACGACACCGCGCGCTCCACCACTGGCTCCAAAAGCAGCAGTA 420
Qy      1474 ArgSerLeuLysGlySerProGlyValThrPheProValHisProLeuAspValMet 1493
Db      421 CGCTCCCTCATCGGAGCGCGCGCGAGCTTCCACCCGTCGACCCCGCTGATGTGATG 480
Qy      1494 AlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGly 1513
Db      481 GCGACGCGCGGACACTGGAACGTGCTGTACGAGGAGAGCTGAAGAGCGCGCCAGGG 540
Qy      1514 ThrAlaSerSerSerGlySerIleAlaArgGlyAlaProValIleValProGluLeu 1533
Db      541 ACCCAGCAAGCTCGGGGGCTCCATTGCGCGCGCGCGCTCATTTGTGCTGAGCTG 600
Qy      1534 GlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHis 1553
Db      601 GGTAAAGCGCGCGAGAACCCCTGACCTATGAGGACCAACGCGGACCCCTTTGCGCGCCAC 660
Qy      1554 LeuProArgGlySerProValThrMetArgGluProThrPro-ArgLeuGlnGluGly-- 1572
Db      661 CTCCACAGAGTTTGGCCGCTGAACCCCGGAGGCCACGCGCGCCCTTGCAGAGGGCA 720
Qy      1573 -----SerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProAr 1590
Db      721 GGCCTTTTCTCCAGCAAGGAGGATCCCGCAGGACCCGAAAGCCTTTGAGCTCCAAACCCCT 780
Qy      1590 gGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProLysSerPr 1610
Db      781 T---CTTGGAAATGGCCCAA-AGTCCCCCGCGGAGCAACCGGGGGGCCCGAAAAACC 836
Qy      1610 o 1610
Db      837 A 837

RESULT 51
BG252161
LOCUS      602365028P1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473271 5',
DEFINITION mRNA sequence.
ACCESSION BG252161
VERSION    BG252161.1 GI:12761977
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10295 row: h column: 08
High quality sequence stop: 711.
Location/Qualifiers
1. .876
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:4473271"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/notes="Organ: liver; vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

ORIGIN

Alignment Scores:

```

Pred. No.:      1,66e-41      Length:      876
Score:          1225.50      Matches:    261
Percent Similarity: 90.24%      Conservative: 7
Best Local Similarity: 87.88%      Mismatches: 24
Query Match:    9.27%      Indels:    9
DB:             12          Gaps:      3

```

US-09-522-753-5 (1-2517) x BG252161 (1-876)

```

Qy      1091 ThrAlaArgProValLeuProArgProThrIleSerAsnProProLeuIleSer 1110
Db      1 ACTGCCGCGCGCTCTGCGCGCGCCACCCACCATCTCCAAACCGCGCTCCCTCATCTCC 60
Qy      1111 SerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSer 1130
Db      61 TCTGCCAAGCACCCAGCGGTCTCTGAGAGGCAATAGTGTGCATCTCCCAAGGATGTCTG 120
Qy      1131 ValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMet 1150
Db      121 GTCCAGCTCCACGTCCTGCTACTCAGAGCATGCCAAGGCGCGGTGGGCGCTGTCAACATG 180
Qy      1151 GlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGlu 1170
Db      181 GGGTGGCCCTGCCCATGAGACCCCAAAAGCTGGACCCCTTCAGCGGAGTGAAGCAGGAG 240
Qy      1171 GlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGln 1190
Db      241 CAGCTGTCCACACGGGCGCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCACAGCCGAG 300
Qy      1191 GluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLys 1210
Db      301 GAGCGCTCCGTGCTGAGAGGAGCAGCTCTGGGCTCAGTTCGGGCGGGAAGCATCAACAAA 360
Qy      1211 GlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThr 1230
Db      361 GGCATTCCAGCACACGGGTGCTCGGAGACGCGCATCATACCCGGGTCTCCATCACC 420
Qy      1231 HisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAsp 1250
Db      421 CACGGCAGCGCAGCTGACGTCTGTACNAGGCGCACCATCACAGGATCATCGCGAGGAC 480
Qy      1251 SerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleThr 1270
Db      481 AGCCCGAGTCTGTCGACCGCGCGGAGGACAGCCTGCCCAAAGGCCACGTCATCTAC 540
Qy      1271 GluGlyLysLysGlyHisValLeuSerTyrGluGlyMetSerValThrGlnCysSer 1290
Db      541 GAAGGCAAGAGGGCCACGCTTGTCTATGAGGGTGGCATGTGTGACCAAGCGTCTCC 600
Qy      1291 LysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThr 1310
Db      601 AAGAGAGCGGACAGAGCAGCTACGACCCCGCATGAGAGCGGCGCCCGGAGCGAC- 658
Qy      1311 TyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeu 1330
Db      659 TATGACATGATGAGGGCGCGGTGGGAGAGCCATCTCTCTCAGC-AGCATCGAAGGTCTC 717
Qy      1331 Met-GlyArgAlaIleProProGluArgHisProHisHisLeuLysGluGlnHisH 1350
Db      718 ATGGGGCGGTGCCATCCCGCGGAGCGGAAA-AAGCCCCACACTACAAAGAGCGGACCCA 776

```

1350 sileArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyr 1370
 Db 777 TCCTCGCG--GTCATCACACAGGATCCCG--GGCCAGTTGAGGCCAGGAGAAC-- 828

1370 rLeuArgGluAlaLysLeuLeuLeuArgGluGlyThrProProPro 1386
 Db 829 ----ACCCGCGTCGGAAGCCAGCCCAAAAGGAGGGCGGCCGCCCA 873

RESULT 52
 CB248833
 LOCUS AL884718.1 763 bp mRNA linear EST 15-JUL-2003
 DEFINITION UI-M-EXO-bv1-k-12-0-UI.r1 NIH BMAP_EX0 Mus musculus CDNA clone IMAGE:5719043 5', mRNA sequence.

ACCESSION CB248833
 VERSION CB248833.1 GI:28387632
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 763)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Straubeberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: pYX-5.

FEATURES
 Location/Qualifiers
 1..763
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5719043"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP_EX0"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGCGGGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,566-41 Length: 763
 Score: 1224.50 Matches: 230
 Percent Similarity: 94.86% Conservative: 10
 Best Local Similarity: 90.91% Mismatches: 12
 Query Match: 9.27% Indels: 1
 DB: 14 Gaps: 1

US-09-522-753-5 (1-2517) x CB248833 (1-763)
 QY 1472 AspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAsp 1491
 Db 3 GAGCTCGCGTCCATCATCGCAGCCGCGCGCTTCCCTGCCCTGCACCGCTGGAC 62
 QY 1492 ValMetAlaAspAlaArgAlaLeuGluAlaCysTyrGluGluSerLeuLysSerArg 1511
 Db 63 ATAATGGCTGACGCCCGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGACCGG 122
 QY 1512 ProGlyThrAlaSerSerGlyGlySerIleAlaArgGlyAlaProValIleValPro 1531
 Db 123 TCAGGAGCAGCAGTGGTGCAGGGGCTCCATCACAGTGGGGCTCCAGTCTGCTGCTCCT 182
 QY 1532 GluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAla 1551
 Db 183 GAACTGGGCAAGCCAGCCCAAGCCCACTGACTTACGAAGACCCAGCGGCACTTCACC 242
 QY 1552 GlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGlu 1571
 Db 243 AGTCACCTGCCAGCTGCCCTGCCCTGTGACCAACGAGGAGCCACGCGCCCTTCAGGAA 302
 QY 1572 GlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGlu 1591
 Db 303 GGCAAGCTCTATCCAGCAAGCGTCCCAAGACCGGAAGCTGACATCTACACCCCGGAG 362
 QY 1592 IleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyr 1611
 Db 363 ATGCCCAAGTCCCAACACAGCACTGTGCCGAGCACCACCTCACCCTCATCCCCCTAT 422
 QY 1612 GluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAla 1631
 Db 423 GAGCACTTGCTCCGGGGCGTGAAGTGTGTGAGCTGTACCGTGTGCATCTACATCCATTGGCC 482
 QY 1632 PheAspProThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyr 1650
 Db 483 TTTGACCCCACTCCATACACCCGAGGATCCCTCTGGAAGCAGCAGCGCGAGCTACTAC 542
 QY 1651 LeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProTyrLeuIle 1670
 Db 543 CTGCCCGGCACTTGGCCCCCAGCCCACTACCCACCTGTACCCACCTTACCTCATC 602
 QY 1671 ArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIle 1690
 Db 603 CGCGGCTACCTCTGACACGGCGGCTTGGAGAACCGCAGACCATCATCAATGACTACATC 662
 QY 1691 ThrSerGlnGlnMetHisHisAsnThrAlaAlaMetAlaGlnArgAlaAspMetLeu 1710
 Db 663 ACCTCGCAGCAGATGACACCAACGCTGCTCCGCCATGCCCCAGCGTGTGACATGCTG 722
 QY 1711 ArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsn 1723
 Db 723 AGGGGTCTGTACCGCGAGAGTCTCGCTGCGCTGCTCAAT 761

RESULT 53
 AL884718/c
 LOCUS AL884718.1 845 bp mRNA linear EST 16-SEP-2002
 DEFINITION AL884718 XGC-egg Silurana tropicalis cdna clone TEG9021004 3', mRNA sequence.
 ACCESSION AL884718.1 GI:22925992
 VERSION AL884718.1
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis

REFERENCE 1 (bases 1 to 845)
 AUTHORS Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
 TITLE Sanger Xenopus tropicalis EST project 2002
 JOURNAL Unpublished (2001)
 COMMENT Contact: Taylor R
 Sanger Institute

ORGANISM
 Silurana tropicalis (western clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Silurana.

Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@anger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS SEQUENCE ID: TEG9021004.q1k7
 Sequencing primer: T7
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 Location/Qualifiers
 1..845
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TEG9021004"
 /dev_stage="egg"
 /lab_host="Escherichia coli XLI-blue"
 /clone_lib="XGC-egg"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dT primed from 5' of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end"

FEATURES source

ORIGIN

Alignment Scores:
 Pred. No.: 3,94e-41 Length: 845
 Score: 1216.00 Matches: 230
 Percent Similarity: 90.78% Conservative: 26
 Best Local Similarity: 81.56% Mismatches: 20
 Query Match: 9.20% Indels: 6
 Gaps: 2

US-09-522-753-5 (1-2517) x AL884718 (1-845)

Qy 195 LysLysGlnGlnLeuGluAlaAlaLysProGluProGluProVal 214
 Db 844 AAAAAACAGCAACAGCTTGAAGAGACGACCAAACTCCAGACCCAGAGCCGCTC 785

Qy 215 SerProProIleGluSerLysHisArgSerLeuValGlnIleTyrAspGluAsn 234
 Db 784 TCCCTCCACCTGTGGAACAGACATCCAGTATAGTCCAAATATTATTATGATGAAAC 725

Qy 235 ArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeu 254
 Db 724 CGGAAAAAGCAGAGGAGCAATAGATCTTAGAAGGTCTAGGTCCAAAAGTAGAGCTG 665

Qy 255 -----ProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsn 269
 Db 664 TCATTTCTTTTGCAGCTCTTTACACACAGCCCTCAGACACAAAGTTTATCATGAAAC 605

Qy 270 IleLysIleAsnGlnAlaMetArgLysLysLeuIleLeuTyrPheLysArgAsnHis 289
 Db 604 ATCAAGACGAATCAGGTGATGAGGAAAAAACTAATCTCTTTAAGAGAAGAAATCAT 545

Qy 290 AlaArgLysGlnTyrLysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeu 309
 Db 544 GCTAGAAAACCTAAGGGAACAGAACATTTGTTCAGCGCTATGATCATGCTGATGGAGCTTGG 485

Qy 310 GluLysLysValGluAlaArgIleGluAsnAsnProArgArgAlaLysGluSerLysVal 329
 Db 484 GAGAGAAGTTGACAGATAGAAAATTAATCTCGAGAAAAGCAAGAGAGCAAAACC 425

Qy 330 ArgGluTyrTyrGluLysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArg 349
 Db 424 CGAGAATATTATGAGAACATTTCTGAAATAGGAAGCAGAGAGCAGCAGCGCT 365

Qy 350 MetGlnSerArgValGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGlu 369
 Db 364 TTCCAG---AGGGTTGGCAGAGAGGAGCTGGGTGTTCAGCAACCATAGCCAGAGTGAA 308

Qy 370 HisGluValSerGluIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMet 389
 Db 307 CATGAATATCTGAATATTGATGGTCTTCTGACAGAGGAATATGAAAAGCAGATG 248

Qy 390 ArgGlnLeuAlaValIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPhe 409

Db 247 CGCAACTCTCAGTCATTCCACCCCATGATGTTTGTATGGGAGCAAGAGGGTAAATTC 188
 Qy 410 IleAsnMetAsnGlyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMet 429
 Db 187 ATTATATGATGTTTGTATGGAGATCCCATGAAGGTGTACAGGACCGGAGTTATG 128

Qy 430 AsnMetTrpSerGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLys 449
 Db 127 AATGTGTGGACTGACCACGAAAGAGATTTTAAAGAAAAATTTGTACACACCTAAA 68

Qy 450 AsnPheGlyLeuIleAlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyr 469
 Db 67 AACTTTGGACTGATAGCCCTCGTATCTGGAAGAAAGAACTGTTTCAGACTGTGTTTATAT 8

Qy 470 TyrTyr 471
 Db 7 TATTAT 2

RESULT 54
 BQ214358
 LOCUS
 DEFINITION AGENCOURT_7594821 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062786
 5', mRNA sequence.
 ACCESSION BQ214358
 VERSION BQ214358.1 GI:20395758
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCDP/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13334 row: n column: 03
 High quality sequence start: 8
 High quality sequence stop: 588.

FEATURES
 Location/Qualifiers
 1..900
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6062786"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.32e-40 Length: 900
 Score: 1204.00 Matches: 255
 Percent Similarity: 90.17% Conservative: 11
 Best Local Similarity: 86.44% Mismatches: 18
 Query Match: 9.11% Indels: 11
 Gaps: 4

US-09-522-753-5 (1-2517) x BQ214358 (1-900)

Qy 234 AsnArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlnValGlu 253

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
plate: LLCM1949 row: o column: 01
High quality sequence start: 13
High quality sequence stop: 527.

```
Location/Qualifiers
1. .1132
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:545452"
   /tissue_type="astrocytoma grade IV, cell line"
   /lab_host="DH10B (phage-resistant)"
   /clone_lib="NIH_MGC_98"
   /notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling H in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit. (Stratagene) and Superscript II RT (Life Technologies note: this is a NIH MGC Library."
```

Scores:		
	3.83e-40	1132
	1195.50	260
Similarity:	73.64%	11
Similarity:	70.65%	60
h:	9.05%	38
	12	8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:
		1132
		260
		11
		60
		38
		8
		Gaps:

151 HisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSer
9 CACCCACAGCAGATCAGCGCACCCCTGCCTCCGCCCCCTCTACTCTCTTCCCTGGCGCCAG
171 CysProValLeuAspLeuArgProProSerAspLeuTyrLeuProProProAspHis
69 TGCCTCCGTCCTGGACCTCCGCGGCCACCCAGTAGTCACTTCACTCCCGCCCTCCGAGCCACCA
191 GlyAlaProAlaArgGlySerProHisSerGluGlyGlyLyAspGlySerProGluProAsp
129 GGTGCCCGCGGCCGTGGCTCCCCCACAGCAGAGGGGCAAGAGGTCTCCAGAGCCACAA
211 LysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGln
189 AAGACGTCCGGTCTTGGGTGGTGGTGTGAGGACGGTATTGAACCTGTCTCCCCACCGAGGGG
231 MetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGln
249 ATGACGGAGGCCAGGGCACTCCCGGAGTGTGTGTACCCCGCTGTCTGTACCGGATGGGGA
251 GlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAla
309 CAGACGGAGCCCGACGAGGATGGGCTCCAAGTCTCCAGGCCAACCCAGCGCAGCGCGCCAG
2271 PhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIle
369 TTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGGTCAAGTCCAGAGAACAGAGATG
2291 AsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGln

TITLE	JOURNAL	COMMENT

FEATURES
SOUND

ORIGIN
Alignment
Pred. No
Score:
Percent S
Best Loc
Query Ma
DB:
US-09-53

QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY

[illegible]

```

Db 429 AACAGAAGCTGAACACCCACCAACCGGAATGAGCTGTAATCAATATACAGCAGCCTCGG 488
Qy 2311 ThrGluilePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyArgSerGln 2330
Db 489 ACGGAGATCTTCAATATGCCCCCATACCGGAACAGGGCTTATGACCTATAGAGCCAG 548
Qy 2331 AlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleAArgLysAlaLeu 2350
Db 549 GCGGTGCAGGAACATGCCAGCAACATGGGCTGGAGGCGCATAAATTAGAAAGGCCCTC 608
Qy 2351 Met-GlyLysTyArgGln-TpGluGluSerProLeuSerAlaAsnAla-PheAsn 2369
Db 609 CTGGGTAAATATGACAGGTGGGAAGATCCCCCGCTCAGCGCCCATGCTTTTACC 668
Qy 2370 ProLeuAsnAlaSerAla---SerLeuProAlaAlaMetProIleThr-AlaAla---- 2386
Db 669 CCTCTGAATGCCAGGGCCCGCTGCGGCTGTATGCCCTTACCCGTGGCTGACG 728
Qy 2387 -AspGlyArgSerAspHisThrLeuThrSerProGlyGly----- 2400
Db 729 GGACGGAGTGGACCAATCACTTCCACTT-CCACCAGCGCGCGCGCGGGAAGGG 787
Qy 2401 -----GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLys 2415
Db 788 CCCAGGTTCTTCTTGGGGAACACCCCGCAGCCCGGAAAGCCCGAGTTCCCCCGGC 847
Qy 2415 sSerProAlaProGlyLeuAlaSerGlyAspArg-----ProProSerVa 2430
Db 848 CCTTCGGGCGCTTGGGGATCTTGGGGGAACCGGCCACCCCTCGCGCCCTTAAGGG 907
Qy 2430 lSerSerValHisSerGluGlyAspGlyAsnArgThrProLeuThrAsn---ArgVa 2449
Db 908 CAACCTCCGAGGAAGACTGGCAACCCCGCCCAACCCGCTTAAACCAACCCCGCGCG 967
Qy 2449 lTTPGluAspArgProSerSer-----AlaGlySerThrProPheProTyArgPro-- 2466
Db 968 GTGGGAGGACCGCCCTCTCTCCCGCGGGGCTCCACCCCGCTTCCCGCTCCCGGTC 1027
Qy 2467 -----LeuIleMetArgLeuGlnAlaGlyValMetAlaSer-- 2478
Db 1028 CCGGGACACAGCTCCCGGACTCGCTTCGGCACCCCGCTTAATCCGCGCTCGGT 1087
Qy 2479 -----ProProPro 2481
Db 1088 GCGGAGCGCCCTCG 1103

RESULT 56
BUI86180
LOCUS
DEFINITION AGENCOURT 7861320 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6141949
5', mRNA sequence.
ACCESSION BUI86180
VERSION BUI86180.1 GI:22700164
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13462 row: p column: 14
High quality sequence stop: 540.

```

```

FEATURES
source
Location/Qualifiers
1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6141949"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 Kb. Library constructed by Life
Technologies."

ORIGIN
Alignment Scores:
Pred. No.: 9,27e-40 Length: 893
Score: 1183.50 Matches: 246
Percent Similarity: 87.94% Conservative: 2
Best Local Similarity: 87.23% Mismatches: 19
Query Match: 8.96% Indels: 15
DB: 13 Gaps: 4

US-09-522-753-5 (1-2517) x BUI86180 (1-893)
Qy 841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGlu 860
Db 6 GAGGAGCAGAGCCCGCGGCTGAGGAGCTGGCAGTGACACAGGAGGCGGAGGAG 65
Qy 861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAla 880
Db 66 CCGCTCAAGCAGAGTGTCACGAGAGAGCCGAGGAGGCGCGCCCAAGGAGACGCG 125
Qy 881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysGlyGlySer 900
Db 126 GAGCGCGCTGAGGCGCAGCGCGGCGGCGCTCAAGGAGAGAGAGAGGCGGAGC 185
Qy 901 GlyArgAlaThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
Db 186 GGCAGGCGCCACACAGCCAGAGCTCGGGCGCCCCCAGGACAGCGACTCCAGTGCCACC 245
Qy 921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuSerPro 940
Db 246 TGCAGTGCAGACGAGGTGGATGAGCGCGGCGCGGAGCAAGACCGGCTGCTGCCCA 305
Qy 941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
Db 306 AGGCGCCAGCTCTCAACCCGACTGGCGACCCCGGCGCAATGCTCTACCCCAAGAGCA 365
Qy 961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaIleProProIleGlnValThrLys 980
Db 366 CTGACCTTGAGCAGCTGAGCAGGAGCGGCTGCCATCCCCCATCCAGGTACACAAA 425
Qy 981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProAlaProPro 1000
Db 426 GTCCATGAGCCCCCGGAGAGCGAGTCCCAAGCAGCTCCCGGAGCGAGCGGAGCGG 485
Qy 1001 ProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
Db 486 CCACCGCAAAACCTGCGAGCGGAGGAGCGGCGGCTCAGCAGCTGGGAGCGAGCGGCGG 545
Qy 1021 GlyLysSerArgSerProAlaProAlaAspLysGlu----- 1033
Db 546 GGCAGAGCAGGAGCGCGGACCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 605
Qy 1034 ---AlaPheAlaAlaGluAlaGlnLysLeuPro-GlyAspProProCysTrpThrSerG 1052
Db 606 CCAGCGCTTGGCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 665
Qy 1052 YLeuProPhe-ProValProProArgGluValIle-LysAlaSerProHisAlaPro--- 1070
Db 666 CTGCGCTTCCCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 725

```

QY 1071 AspProSerAlaPheSer-TyrAlaProProGlyHisProLeu---ProLeuGlyLeuHi 1089
 Db 726 CCCCCTCAAGCTTTCTCCCTAGCTCCCTCCCTGGGTGCTCCACCGGCCCTCCCTCC 785
 QY 1089 sAspThrAlaArgProValLeuPro---ArgProProThrIleSerAsnProProPro 1107
 Db 786 ATGAACACTGGCCCGCCGGTCTCTGGCGGGGCCCAACCCACCATTTCTCCAAACCG 843

RESULT 57
 BU257626 790 bp mRNA linear EST 26-NOV-2002
 LOCUS 603414849F1 CSEQHN38 Gallus gallus cdna clone CHEST337b13 5', mRNA
 DEFINITION sequence.

ACCESSION BU257626
 VERSION BU257626.1 GI:25518285
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 790)
 AUTHORS Boardman,P.E., Sanz-Exquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
 TITLE A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392

COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1..790
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, HiseX"
 /db_xref="taxon:9031"
 /clone="CHEST337b13"
 /dev_stage="36"
 /lab_host="DH10B"
 /clone_lib="CSEQHN38"
 /note="Organ: limbs; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Alignment Scores:
 Pred. No.: 8.5e-40 Length: 790
 Score: 1183.00 Matches: 243
 Percent Similarity: 89.89% Conservative: 6
 Best Local Similarity: 87.73% Mismatches: 8
 Query Match: 8.95% Indels: 21
 DB: 13 Gaps: 3

US-09-522-753-5 (1-2517) x BU257626 (1-790)
 QY 259 GlnProSerAspThrArgGlnTyrHisGluAsnIleTysIleAsnGlnAlaMetArgLys 278

Db 2 CAGCCTTCAGACACCCAGGCAATACCATGAAACATTAAATAACACAGGCAATGCGGAAG 61
 QY 279 LysLeuIleLeuTyrPheLysArgAsnHisAlaArgLysGlnTyrLysGlnLysPhe 298
 Db 62 AAGTTAAATTTTGTACTTTAAGAGAGAAATCATGCTCGCAAGCAGTGGGAAACAGAGTTTC 121
 QY 299 CysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsn 318
 Db 122 TGCACGGCTATGACCCAGCTGATGGAGGCTGGGAAAGAAAGTGGAGCGCATTTGAAAT 181
 QY 319 AsnProArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPhePro 338
 Db 182 AACCCCGCGCGCTGCCAAGAGAGCAAAAGTCGGGAATACTATGAGAAACAAATTCCT 241
 QY 339 GluIleArgLysGlnArgGluLeuGlnGluArgMetGln-----SerArgVal 354
 Db 242 GAGATCCGAAGCAGCGGGAACCTGCAAGACCGCATGCAGAGGTAGAGGAGACCGGTA 301
 QY 355 GlyGlnArgLysSer---GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSer 373
 Db 302 GGACAGAGGGGCGAGCGGAGGCTCTCCATGTCGCTGCGCAGTGAACACAGAGTGTGC 361
 QY 374 GluIleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAla 393
 Db 362 GAGATCATCGATGGCTCTCAGAGCAAGAGAACCTGGAGAAGCAGATGCCCGCAGCTGCC 421
 QY 394 ValIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsn 413
 Db 422 GTATCCACCCATGCTCTATGATGTCACAGCAGCGCATTAATTCATCACTCACTGAC 481
 QY 414 GlyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTyrSer 433
 Db 482 GGCTCATGACGACCCCATGAAGTCTACAAGACCGCAGGTGATGAACATGTGGAGC 541
 QY 434 GluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeu 453
 Db 542 GAGCAGAGAGAGGAGACCTTCCGGGAAAAGTTTCATGACAGCACCACCAAGAACTTTGGCGT 601
 QY 454 IleAlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThr 473
 Db 602 ATTGCTCTCTTTCTGGACAGAGAGACGCTGGCAGACTGCTCTGTACTACTACTGACC 661
 QY 474 LysLysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArgArg--GlyLys 493
 Db 662 AAGAAGAATCAGAACTACAAGAACCTCGTGAGGAGGAATTACCGCGCGGTGGGAAAAA 721
 QY 493 erGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetP 513
 Db 722 GCCAGCAGCAGCAAA-----ATGC 742
 QY 513 roArgSerSerGlnGluLysAspGluLysGluLysGluLysGluLysGlu 528
 Db 743 CCGG-AGCAGACAGAGAGAGAACGATGAACAGAGAGAGAGAGAGAG 788

RESULT 58

LOCUS BQ070408

DEFINITION AGENCOURT 6839381 NIH_MGC_47 Homo sapiens cdna clone IMAGE:5922788

5', mRNA sequence.

ACCESSION BQ070408

VERSION BQ070408.1 GI:19899454

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1032)

NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgaaps-remail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
High quality sequence stop: 623.
Location/Qualifiers
1. 1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5922788"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_4"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 2.44e-39 Length: 1032
Score: 1175.00 Matches: 246
Percent Similarity: 89.75% Conservative: 8
Best Local Similarity: 86.93% Mismatches: 19
Query Match: 8.89% Indels: 10
DB: 13 Gaps: 5

US-09-522-753-5 (1-2517) x BQ070408 (1-1032)

Qy 1 MetSerGlySerThrGlnLeuValAlaGlnThrTTPArgAlaThrGluProArgTyrPro 20
Db 109 ATGTCGGGCTCCACACGCTGTGGCAGACGTGGAGGGCCATGAGCCCGCTACCCG 168
Qy 21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
Db 169 CCCACACGCTTCTTACCAGTCAGATGCCCGGACGACACGAGCTCGGGCTCTGT 228
Qy 41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleGln 60
Db 229 GAGTACCAGCACCACTCCCGGACTATGCTCCACCTGCCCGGCTCCCATCATCCAG 288
Qy 61 ProGlnArgArgProSerLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
Db 289 CCCAGCGCGGAGCGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCAG 348
Qy 81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
Db 349 GAGCTCCACCTGGCGGACAGATCCCACTATACCTGCCCGAGTGGGGAATCAGAGATG 408
Qy 101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
Db 409 GAGTTTCATTGAAGACGAGCGCCCTCGCTAGAGTGTGCTGACCCCTGCTGCGACCG 468
Qy 121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
Db 469 TCACCCCTGCTGGCCACGGGCGAGCTCGGGATCTGAAGACCTCACCAAGGACCGTAGC 528
Qy 141 LeuThrGlyLysLeuGluProValSerProProSerProHisThrAspProGluLeu 160
Db 529 CTGACGGGCAGCTGGAAACCGGTGTCTCCCGCCAGCCCCCGCACACTGACCTTGAGCTG 588
Qy 161 GluLeuValProProArgLeuSerLysGluLeuIleGlnAsnMetAspArgValAsp 180

Db 589 GAGTGGTGGCCGCCACGGCTGTCCAAGGAGGAGCTGTATCCAGAACATGACCGCGTGGAC 648
Qy 181 ArgGluIleThrMetValGluGlnIleSerLysLeuLysLysGlnGlnGlnLeu 200
Db 649 CGAGGATCACCATGGTAGAGCAGCAGATCTCTAGCTGAAGAGAACACGACGACGCTG 708
Qy 201 GluGluGluAlaLysProGluProGluLysProValSerProProProIleGlu 220
Db 709 GAGGAAGAGGCTGCCAAGCCGCCGAGCTGAGAAACCGTGTACCCGCCGCCATCGAG 768
Qy 221 SerLysHis---ArgSerLeuValGlnIleIleTyrAspGluAsn-ArgLysLysAla-G 239
Db 769 TCGAAGCACCCCGCCCTGGTGGTCAAAATCTTCTCCACCGAGAACCCCGAAAAGGCTTG 828
Qy 239 luAlaAlaHisArgIleLeu---GluGlyLeuGly---ProGlnValGluLeuProLeu- 256
Db 829 AAGCTGGACATCGGAATTCGGGAAGGCTTGGGGGCCCCAGGGGGAACCTGCCGCTG 888
Qy 257 --TyrAsnGln--ProSerAspThrArg-----GlnTyrHisGluAsnIleLysIleAs 273
Db 889 GTACAAACACGACCCCTCCGAAACCCGGCGGTAATCATGAAGAACCTCTAAATAATAA 948
Qy 273 n 273
Db 949 C 949
RESULT 59
BX732682/C
LOCUS BX732682
DEFINITION BX732682 XGC-tadpole Silurana tropicalis cDNA clone TTPA071k07 3', mRNA sequence.
ACCESSION BX732682
VERSION BX732682.1 GI:38405423
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 858)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
UNPUBLISHED (2003)
CONTACT: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE ID: TTPA071k07.g1kat7
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Nigel Garrett.
cDNA was oligo dr primed from 5ug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1. 858
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA071k07"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dr primed from 5ug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Alignment Scores:

```
Pred. No.: 7.23e-39 Length: 858
Score: 1161.50 Matches: 223
Percent Similarity: 86.38% Conservatives: 37
Best Local Similarity: 74.09% Mismatches: 24
Query Match: 8.79% Indels: 17
DB: 13 Gaps: 3

US-09-522-753-5 (1-2517) x BX732682 (1-858)

Qy 232 AspGluAsnArgLysLysLysAlaGluAlaHisArgIleLeuGluGlyLeuGlyProGln 251
Db 858 GATCAAAACCGGAAAAAGCAGAGAACATAGATCTTAGAGGTCTAGGTCCAAAA 799
Qy 252 ValGluLeuProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLys 271
Db 798 GTAGAGCTGCCCTCTTTACACCCCTTCAGACACAAAGATTATCATGAAACATCAAG 739
Qy 272 IleAsnGlnAlaMetArgLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArg 291
Db 738 ACGAATCAGGTGATGAGGAAAAAATAACTACTGTTCTTAAAGAGAGAAATCATGTAGA 679
Qy 292 LysGlnTyrLysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLys 311
Db 678 AAATTAAGGGAACAGAAACATTTGTTCAGCGCTATCATGCTGATGCGAGGCTTGGGAG 619
Qy 312 LysValGluArgIleGluAsnProArgArgAlaLysGluSerLysValArgGlu 331
Db 618 AAGTTGACAGATAGAAAATAATATCTCGGAGAAAGCAAGAGAGCAAAACCCGAGAA 559
Qy 332 TyrTyrGluLysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGln 351
Db 558 TATTATGAGAGCAATTTCTCTGAAATAGGAAGCAGAGAGAGCAGCAGGAGCGTTCCAG 499
Qy 352 SerArgValGlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGlu 371
Db 498 ---AGGGTTGGCGAGAGAGCTGGCTGTCAGCAACCATAGCCAGGAGTGAACATGAA 442
Qy 372 ValSerGluIleIleAspGlyLeuSerGluGlnGluAsnLeuLysGlnMetArgGln 391
Db 441 ATAICTGAATATTATGATGCTTCTTCTGAAACAGAGAGATATGAAGAAGCAGATCGGCCAA 382
Qy 392 LeuAlaValIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsn 411
Db 381 CTCTCAGTCATTCCACCCATGATGTTGATCGGAGCAAGAGCGGTAAAAATTCATTAAT 322
Qy 412 MetAsnGlyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMet 431
Db 321 ATGAATGTTTGTATGAGAGATCCCATGAAAGGTGTACAAAGGACCGGCGATTTATGAATGTG 262
Qy 432 TrpSerGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPhe 451
Db 261 TGGAGTACCACGAAAGAGATTTTAAAGAAAATTTGTACACACCCCTAAAACTTT 202
Qy 452 GlyLeuIleAlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyr 471
Db 201 GGACTGATAGCTCGTATCTGAAAGAAAGACTGTTTCAGACTGTGTTTTATATTAT 142
Qy 472 LeuThrLysLysAsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArg 490
Db 141 TTAACCTAAGAAAAATGAAATTTTAAAGCCCTTGTGAGAGAAACTACCTCTAAAGAGGA 82
Qy 491 GlyLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 510
Db 81 GGACCAACACGACACAG-----64
Qy 511 ProMetProArgSerSerGlnGluLysAspGluLysGluLysGluLysGluLysGlu 530
Db 63 ---ATAACCCGCTGCCCAAGAGAAAGGAAATGAAATGAAATGAGAGAAAAA 7
Qy 531 Lys 531
Db 6 AAA 4
```

RESULT 60
BE793487
LOCUS
DEFINITION
787 bp mRNA linear EST 20-SEP-2000
G01588814F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3943059 5',
mRNA sequence.
BE793487
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 787)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1C0797 row: 1 column: 04
High quality sequence stop: 762.
FEATURES
Location/Qualifiers
1..787
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3943059"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/notes="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

Alignment Scores:
Pred. No.: 8.75e-39 Length: 787
Score: 1158.50 Matches: 246
Percent Similarity: 96.11% Conservatives: 1
Best Local Similarity: 95.72% Mismatches: 8
Query Match: 8.77% Indels: 7
DB: 10 Gaps: 1

US-09-522-753-5 (1-2517) x BE793487 (1-787)

Qy 2215 LeuGlyGlyGlyGluAspGlyIleGluProValSerProGluGlyMetThrGluPro 2234
Db 4 CTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCACCGAGGGCATGACGAGCCA 63
Qy 2235 GlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluPro 2254
Db 64 GG-CACCTCCCGAGTGTGTGTACCCGCTGTGTGTACCGGGATGGGAACACGAGAGCCC 122
Qy 2255 SerArgMetGlySerLysSerProGlyAsnThrSerGlnProAlaPhe-PheSerLys 2274
Db 123 AGCAGATGGGCTCCAGTCTCCAGGACACACAGCAGCAGCGCCGCTCTTACAGCAA 182
Qy 2274 sLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluLeuAsnLysLysLe 2294
Db 183 GCTGACCGAGAGCAACTCCGCGCATGGTCAAGTCCCAAGAACAGAGATCAACAAGAGCT 242
Qy 2294 uAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePh 2314

```

Db      243 GAACACCCCAACCGGAATGACCTGAATACAAATATCAGCCAGCCTGGAGCGAGATCTT 302
Qy      2314 eAenMetProAlaIleThrGlyLeuMetThrTyrArgSerGlnAlaValGlnG 2334
Db      303 CAATATGCCCGCATCACCGGAACAGGCTTATGACCTATAGAACCGAGCGGTGACGA 362
Qy      2334 uHisAlaSerThrAsnMetGlyLeuAlaIleIleArgLysAlaLeuMetGlyLys 2354
Db      363 ACATGCCAGCACCAACATGGGCTGGAGGCCATAATTAGAAAGGCACATCATGGTAAATA 422
Qy      2354 rAspGlnTrpGluGluSerProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSe 2374
Db      423 TGACAGATGGGAAGAGTCCCGCCCTCAGCCCAATGCTTTAAACCTCTGAATGCCAG 482
Qy      2374 rAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLe 2394
Db      483 TGCCAGCTGCCCGCTGTATGCCATTAACCGCTCTGACGAGCGAGGTGACACACACT 542
Qy      2394 uThrSerProGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArgLysAl 2414
Db      543 CACCTCGCCAGGTGGCGGGGGAAGGCAAGGTCTCTGCGACACCCAGCAGCGGAAAGC 602
Qy      2414 aLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSerValHi 2434
Db      603 CAAGTCCCGCGCCCGGCTGGCATCTGGGACCGGCA-CCTCTGTCTCTCTCAGTGCA 661
Qy      2434 sSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAspArgPr 2454
Db      662 CTCGGAGGAGACTCAACCG-CGAGCCCGTC-ACCAACCGCGTGGGAGCAGCGCC 719
Qy      2454 oSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArg 2470
Db      720 CTCGTCCGAG--TTACGCGCATTCCT--AAATCCCTGATCATGCGG 763

RESULT 61
CA749602
LOCUS      718 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-FY0-cde-a-15-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6831736 5', mRNA sequence.
ACCESSION CA749602.1 GI:25571972
VERSION CA749602.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5,
Location/Qualifiers
1..718
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6831736"
/tissue_type="whole brain"
/dev_stages="embryo 13.5,14.5,16.5,17.5dpf"
/lab_host="DH10B (T1 phage resistant)"

```

/clone lib="NIH_BMAP_FY0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to the protocol of R. B. Doolittle, R. B. Doolittle,
 Ronald D. Lennan and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is ACCGAGACAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

```

Alignment Scores:
Pred. No.: 1,27e-38 Length: 718
Score: 1153.50 Matches: 216
Percent Similarity: 94.14% Conservative: 9
Best Local Similarity: 90.38% Mismatches: 13
Query Match: 8.73% Indels: 1
DB: 14 Gaps: 1
US-09-522-753-5 (1-2517) x CA749602 (1-718)

Qy      1480 ProGlyArgThrPheProValHisProLeuAspValMetAlaAspAlaArgAlaLeu 1499
Db      2 CCCGCCGCGCTTCCCTCGCCCTGCACCGCTGGACATAATGGCTGAGCGCGGCACTG 61
Qy      1500 GluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerGly 1519
Db      62 GAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCGGTCAGGAGCCAGCAGTGGTGAGG 121
Qy      1520 GlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSer 1539
Db      122 GGCTCCATCACACGTTGGGCTCCAGTCTGTCGCTGAACCTGGGCAAGCCACGCAAGC 181
Qy      1540 ProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerPro 1559
Db      182 CCACCTGACTTACGAAGACACCGCGGACCCCTTCCACAGTCACCTGCCAGTGGCTCCCT 241
Qy      1560 ValThrMetArgGluProThrProArgLeuGluGlySerLeuSerSerLysAla 1579
Db      242 GTGACACGAGGGAGCCCGCCCGCTTCCAGGAAGGAGCGCTCTTATCCAGCAGGCGG 301
Qy      1580 SerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThr 1599
Db      302 TCCCGAGGACCGGAGCTGACATCTACACCCCGGAGATCGCAAGTCCCCACACGACT 361
Qy      1600 ValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSer 1619
Db      362 GTGCCCGAGCAGCACCCCTCACCCCTATGAGCAGCTTCTCCGGGGCGTGACT 421
Qy      1620 GlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArg 1639
Db      422 GGTGTGAGACCTGTACCGGGTGCATATCCATTTGGCTTTTGACCCCGCCACCTCCATCC 481
Qy      1640 GlyIleProLeuAsp---AlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsn 1658
Db      482 GGGATCCCTCTGGAGAGCAGCAGCGCCCTACTACCTGCCCGGCACTTGGCCCCCAGC 541
Qy      1659 ProThrTyrProHisLeuTyrProProTyrIleuIleArgGlyTyrProAspThrAlaAla 1678
Db      542 CCCACCTACCCACACCTGTACCCACCTTACCTCATCCCGCGGCTACCTGACACGCGGCG 601
Qy      1679 LeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHis 1698
Db      602 CTGGAGAACCCCGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGACACCAAC 661
Qy      1699 ThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGlu 1717

```


/db_xref="taxon:9606"
 /clone="IMAGE:383162"
 /tissue="type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 20"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACAGG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.: 2,4e-37 Length: 804
 Score: 1124.00 Matches: 243
 Percent Similarity: 88.21% Conservatives: 4
 Best Local Similarity: 86.79% Mismatches: 22
 Query Match: 8.51% Indels: 16
 DB: 10 Gaps: 1

US-09-522-753-5 (1-2517) x BE728145 (1-804)

QY 81 GluLeuHisLeuArgProGluSerHisSerTyLeuProGluLeuGlySerGluMet 100
 |||||
 Db 1 GAGCTCCACCTCGCGCCAGAGTCCACCTCATCTCTCCGAGCTGGGAAAGTCAGAGATG 60
 QY 101 GluPheLeuGluSerLeuArgProArgLeuGluLeuProAspProLeuLeuArgPro 120
 |||||
 Db 61 GAGTTTCATTAAGAAAGACGCGCTCGCTAGAGCTGCTGCTGAGCCCTGCTGCGACCG 120
 QY 121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLeuAspArgSer 140
 |||||
 Db 121 TCACCCCTGCTGCGCCAGCGCGCTGCGGATCTGAAGCTTCCACAGGACCGCTAGC 180
 QY 141 LeuThrGlyLeuLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
 |||||
 Db 181 CTGACGGGAGAGCTGGAACCGGTGTCTCCCTCCAGCCCGCCACACCTGACCTGAGCTG 240
 QY 161 GluLeuValProProArgLeuSerLeuGluGluLeuLeuGlnAsnMetAspArgValAsp 180
 |||||
 Db 241 GAGCTGGTGGCCCGCCAGCGGTGTCCAGGAGGAGAGCTGATCCAGAACATGCGCGCTGGAC 300
 QY 181 ArgGluIleThrMetValGluGlnGlnIleSerLeuLeuLeuLysLysGlnGlnLeu 200
 |||||
 Db 301 CGAGAGATCACCATTGGTAGAGCAGAGATCTTAAGCTGAAGAGAGCAGCAAGCAGCTG 360
 QY 201 GluGluGluAlaAlaLeuProGluProGluLysProValSerProProProIleGlu 220
 |||||
 Db 361 GAGGAGGAGGCTGCCAAGCGCGGAG-CCTGAGAAGCCCGTGTACCCGCGCCATC-GAG 418
 QY 221 SerLysHisArgSerLeuValGlnIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 240
 |||||
 Db 419 TCGAAGCAGCGCAGCGCTGGTGCAGATCATCTACGACGAGAACCGGAAAGAGGCTGAAGCT 478
 QY 241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyraGlnPro 260
 |||||
 Db 479 GCACATCGATTCTGGAAGGCTCGGGCCCGAGGTGGAGCTGCCCTGTACACCGAGCC 538
 QY 261 SerAspThrArgGlnTyraHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
 |||||
 Db 539 TCCGACACCGCGCAGTATCATGAGAATCAATCAAAATAACACGCGCATCGGAGAGGCTA 598
 QY 281 IleLeuTyraPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
 |||||
 Db 599 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGAAACCAATGGGAGCAGAGATC-TGCCAG 657
 QY 301 ArgTyraAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
 |||||
 Db 658 CGCTATGACCAAGTC-ATGAGGCGCTGGGAGGACGAAGGTGGAGCGCATCGAGAACACCCC 716

QY 321 ArgArgArgAlaLysGluSerLysValArgGluTyraTyraGluLysGlnPheProGluLeu 340
 |||||
 Db 717 CGGCGGGG-GCCAGGAGACGACGAGGGGGGAGTAATA----- 751
 QY 341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
 |||||
 Db 752 -----CGAGAGCAGGTCTGTGAGAACGCGAGCGCGGAGCGATCGAGGGGG 802
 RESULT 64
 LOCUS BM944466
 DEFINITION UI-M-EH0p-bvr-k-08-0-UI.r1 NIH-BMAP_EH0p Mus musculus cDNA clone
 IMAGE:5695999 5', mRNA sequence.
 ACCESSION BM944466
 VERSION
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 779)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

Location/Qualifiers
 1..779
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5695999"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH-BMAP_EH0p"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is CAGCCAGCAG. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institute of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 2,55e-37 Length: 779
 Score: 1123.00 Matches: 225
 Percent Similarity: 88.89% Conservatives: 7
 Best Local Similarity: 86.21% Mismatches: 26
 Query Match: 8.50% Indels: 5
 DB: 12 Gaps: 2

US-09-522-753-5 (1-2517) x BM944466 (1-779)

```
Qy 2079 LysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGly 2098
Db 2 GAAGAGCACTTGGAGGGAGCTGGGCACAGCAGCAGCGCCCATGAGACTCAGCGCG 61
Qy 2099 GluAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerPro 2118
Db 62 GAGGTGCGCCATCTCCACATCTGGGCCACTGCGCCAGAGCAGCGCCCTCATCTAGCCCA 121
Qy 2119 LeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHis 2138
Db 122 CTCTCTCAGAGTGGCCCGCATCAAGAGTCAACAGAGGGTGGTCACTTGGCTCAGCAC 181
Qy 2139 IleSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaPro 2158
Db 182 ATCAGTAGGTCATTACGACAGACTACACGCCGCCACACCCGCGAGCGCTCAGTGGGCC 241
Qy 2159 LeuProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArg 2178
Db 242 CTTCGCCGCCCTCTCTACTCTTCCCGAGCGCAGCTGCGCTGTCTGGATCTTCGCCGC 301
Qy 2179 ProProSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerPro 2198
Db 302 CCACCCAGTACCTCTACTCTCCACCCCGCCAGCATGCGCACCCGCGCGGATCCGCC 361
Qy 2199 HisSerGluGlyGlyLysArgSerProGluProAlaSerLysThrSerValLeuGlyGly 2218
Db 362 CACAGTAGGGGGGCAAGGTCTCCAGACCCAGCAAAACATCGTCTGGCGAGCAGC 421
Qy 2219 GluAspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArg 2238
Db 422 GAGGATGCCATTGAGCCTGTGTCCCCACAGAGGCGATGACTGAGCCAGGACATGCTCG 481
Qy 2239 SerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGly 2258
Db 482 AGCACTGGGTACCCACTCTGTATCGAGCGGGGAACAGCGCGGCC---AGGATGGGC 538
Qy 2259 SerLysSerProGlyAsnThrSerGlnProProAlaPhe-PheSerLysLeuThrGluSe 2278
Db 539 TCTAAGTCTCCAGGCACACAGCAGCGCGCGCGCTTCTTTCAGTAGCTGACTGAGAG 598
Qy 2278 rAsnSerAlaMetValLysSerLysGlnGluIleAsnLysLysLeuAsnThrHisAs 2298
Db 599 CAC-TCCGCCATGGTGAAGTCGAAGCAGCAGAGATCAACANGAACTCAACACCCACA 657
Qy 2298 nArgAnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAl 2318
Db 658 CCGGAACGAGCGAGAAATCAATATTGGCGCGCTGGGACGGAATCTTCANNCATGCCGCC 717
Qy 2318 atIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerTh 2338
Db 718 -ATCACTGGAGCAGGCTTATGNCCTGTAGAGCGCGGTGCAGA---CAGCGCAGCAC 773
Qy 2338 r 2338
Db 774 A 774

RESULT 65
BX771375/c
LOCUS
DEFINITION BX771375 XGC-egg Silurana tropicalis cDNA clone Tegg036f09 3', mRNA
ACCESSION BX771375
VERSION BX771375.1 GI:39678580
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
REFERENCE
1 (bases 1 to 811)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
TITLE
```

Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxtom, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: Tegg036f09.q1k17
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dT primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5', end and NotI at the 3', end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1..811
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="Tegg036f09"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

Alignment Scores:
Pred. No.: 5,2e-37 Length: 811
Score: 1116.00 Matches: 216
Percent Similarity: 88.93% Conservative: 25
Best Local Similarity: 79.70% Mismatches: 23
Query Match: 8.44% Indels: 7
DB: 13 Gaps: 2

US-09-522-753-5 (1-2517) x BX771375 (1-811)

Qy 204 AlaAlaLysPro-ProGluProGluLysProValSerProProLleGluSerLysHi 223
Db 810 GCAGCAAAACCTCCAGAGCCAGAGAGCCGCTCCCTCCCTCCACCTGGGAACAGAGCA 751
Qy 223 sArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaHisAr 243
Db 750 TCGCAGTATAGTCAATATTATGATGATAAAGCGGAAAGACAGAGGAGCACATAA 691
Qy 243 gIleLeuGluGlyLeuGlyProGlnValGluLeu-----ProLeuTyrAs 258
Db 690 GATCTAGAAAGTCTAGGTCCAAAGTAGAGCTGTCTATTTTTCAGCGCTCTTTTCAA 631
Qy 258 nGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLy 278
Db 630 CCAGCCCTCAGACACAAAGTTTATCATGAAAAATCAAGACGAATCAGGTGATGAGGAA 571
Qy 278 sLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTyrLysPh 298
Db 570 AAAACTAATACTGTCTTTAAGAGAGAATCATCATGAGAACTAAGGGAACAGAACAT 511
Qy 298 eCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAs 318
Db 510 TTGTTCAGCGCTATGATCAGCTGATGGAGGCTTGGGAGAAAGTTGACAGGATAGAAAA 451
Qy 318 nAsnProArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPhePr 338
Db 450 TAATCTCTCGGAGAAAGCAAGAGGAGAGCAAAACCCGAGATATTTATGAGAGCAATTTCC 391
Qy 338 oGluIleArgLysGlnArgGluLeuGlnGlnArgMetGlnSerArgValGlyGlnArgGl 358
Db 390 TGAATAAGGAAGCAGAGAGAGAGAGAGAGAGCGGCTTTCCAG---AGGGTTGGCGCAGAGAG 334
Qy 358 ySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGl 378


```

VERSION BM783748.1 GI:19131980
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 651)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
UNPUBLISHED (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: A column: 06
High quality sequence stop: 651.
Location/Qualifiers
1..651
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SSSNU484-36-A06"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10P"
/clone_lib="SSSNU484"
note="Organ: Stomach; Vector: pTZ18P1; Site:1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transfection of
competent cells E. coli Top10P by electroporation
method."
ORIGIN
Alignment Scores:
Pred. No.: 6,578-37 Length: 651
Score: 1111.00 Matches: 215
Percent Similarity: 81.75% Conservative: 0
Best Local Similarity: 81.75% Mismatches: 2
Query Match: 8.41% Indels: 46
DB: 12 Gaps: 1
US-09-522-753-5 (1-2517) x BM783748 (1-651)
QY 2145 GlnAspTyrThrArgHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyr 2164
Db 1 CAGGACTACACCGGCACCCACACAGAGCTCAGCGCACCTCGCCGCCGCCCTCTAC 60
QY 2165 SerPheProGlyAlaSerCysProValLeuAspLeuArgArgProSerAspLeuTyr 2184
Db 61 TCTTCTCCCTGGGGCAGCTGCCCCCTCTGGACCTCCCGCCGCCACCCAGTGACCTCTAC 120
QY 2185 LeuProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLys 2204
Db 121 CTCGCCGCCCGACCATGGTCCCGCCCGCGCTCTCCGCCACAGCGAGGGGGCAAG 180
QY 2205 ArgSerProGluProAsnLysThrSerValLeuGlyGlyGluAspGlyIleGluPro 2224
Db 181 AGGTCTCCAGAGCCAAACAGACGTCGGTCTTGGGTGGTGGTGGAGGCGGTATTGAACCT 240

```

2225 ValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeu 2244
 241 GTGTCCCCACCGAGGGCATGACGGAGCCAGGCACTCCCGAGTGTGTATCCCGCTG 300
 2245 LeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsn 2264
 301 CTGTACCGGATGGGGAACAGACGAGCCAGCAGGATGGCTCCCAAGTCTCCAGGCAAC 360
 2265 ThrSerGlnProAlaPhePheSerLysLeuThrClnSerHsnSerAlaMetValLys 2284
 361 ACCAGCCAGCGCCAGCTTCTTCAGCAAGCTACCGAGAGCACTCCGCCATGGTCAAG 420
 2285 SerLysLysGlnGluLeuAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyr 2304
 421 TCCAGAGAGCAAGATCAACAGAGAGCTGAACACCCACACCCGATGAGCGCTGATAC 480
 2305 AsnIleSerGlnProGlyThrGluLeuPheAsnMetProAlaIleThrGlyLysLeu 2324
 481 AATATCAGCCAGCTGGGAGCGAGATCTTCAATATGCCGCCATCACCGGAACAGGCTT 540
 2325 MetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLysGluAla 2344
 541 ATGACCTATTAGAGCCAGCGGTCAGAGAACATGCCAGACCAACATGGGCTGGAGGC 600
 2345 IleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSer 2364
 601 CATATTAGAAAGGCATCTCATG----- 621
 2365 AlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThr 2384
 621 ----- 621
 2385 AlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLys 2404
 622 -----GTTGGCGCGGGAAGGCCCAAG 642
 2405 ValSerGly 2407
 643 GTCTCTGCC 651

RESULT 68
 CF540568
 LOCUS IMAGE:30545297 5', mRNA sequence.
 DEFINITION U1-M-GVO-cie-e-18-0-UI.r1 NIH_BMAP_GVO Mus musculus cDNA clone
 CF540568
 VERSION CF540568.1 GI:34593091
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 683)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: csapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..683
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"

FEATURES
 source

/db_xref="taxon:10090"
 /clone="IMAGE:30545297"
 /tissue_type="whole brain"
 /dev stage="1.5, and 15 days newborn"
 /lab host="DH10B (T1 phage resistant)"
 /clone lib="NIH BMAP GV0"
 /note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag is CGAAGTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 7.27e-37 Length: 683
 Score: 1110.50 Matches: 210
 Percent Similarity: 94.71% Conservative: 5
 Best Local Similarity: 92.51% Mismatches: 11
 Query Match: 8.40% Indels: 1
 DB: 14 Gaps: 1

US-09-522-753-5 (1-2517) x CP540568 (1-683)

Qy 1532 GluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAla 1551
 Db 1 GAATGGGCAAGCCAGCGCAAGCCACTGACTACGAGAGCCAGCGGCGACCTTCACC 60
 Qy 1552 GlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGlu 1571
 Db 61 AGTCACCTGCCAGCTGGTCCCTGTGACCGAGGAGGCCACGCCACGCTTCAGGAA 120
 Qy 1572 GlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGlu 1591
 Db 121 GGCAGCTTCCTATCCAGACGCGTCCAGACCGGAGAGCTGACATCTACACCCCGGAG 180
 Qy 1592 IleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyr 1611
 Db 181 ATGCCAAGTCCCAACAGCACTGTGCCGAGACCACTTCACCCCATCTCCCTAT 240
 Qy 1612 GluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAla 1631
 Db 241 GAGCACTTGCTCCGGGGGTGACTGGTGTGACCTGTACCGTGGTGCACATCCCATGGCC 300
 Qy 1632 PheAspProThrSerIleProArgGlyIleProLeuAsp----AlaAlaAlaTyrTyr 1650
 Db 301 TTTCACCCACCTCCATACCCCGAGGATCCCTCTGGAAGCAGCAGCGCGCTACTAC 360
 Qy 1651 LeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrIleuile 1670
 Db 361 CTGCCCCGGCACTTGCCCCCAGCCCCACCTTACCACACCTGTACCCACCTTACCTCATC 420
 Qy 1671 ArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIle 1690
 Db 421 CGCGGTACCTTGACACCGCGGCTTGAGAACCGCCAGCATCATCATGACTACATC 480
 Qy 1691 ThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeu 1710
 Db 481 ACCTCGCAGCAGATGCACCAACGCTGCCTCCGCCATGGCCCGCGTGCATGCTG 540
 Qy 1711 ArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGly 1730
 Db 541 AGGGGTCTGTACCGCGAGAGTCTCGCTGCGCTGCGCTCAATTATGCGCTNGCCGAGGCG 600

Qy 1731 IleIleAspLeuSerGlnValProHisLeuProValLeuValProThrProGlyThr 1750
 Db 601 ATTATCGACCTGTCCCAAGTCCACACCTGCCGTGCTGCTGCCCAACAGCGGCGAC 660
 Qy 1751 ProAlaThrAlaMetAspArg 1757
 Db 661 CTGCCACCGCCATCGACCGC 681
 RESULT 69
 AI830862
 LOCUS
 DEFINITION
 wj1b09.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2407289 3,
 similar to TR:000613 O00613 T3 RECEPTOR-ASSOCIATING COFACTOR-1. [1]
 ; mRNA sequence.
 ACCESSION
 AI830862
 VERSION
 AI830862.1 GI:5451533
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 622)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL
 COMMENT
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 2006 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 465.
 Location/Qualifiers
 1..622
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2407289"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev stage="adult"
 /lab host="DH10B (phage-resistant)"
 /clone lib="NCI CGAP Lul9"
 /note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 p7T73 vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

ORIGIN

Alignment Scores:
 Pred. No.: 9.15e-37 Length: 622
 Score: 1107.00 Matches: 207
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.38% Indels: 0
 DB: 9 Gaps: 0

US-09-522-753-5 (1-2517) x AI830862 (1-622)

Qy 2117 SerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgValThrLeuAla 2136

```

Db      2 AGCCCGCTGCTCCAGACCGCCCGGAGGTCAAGAGTCAACAGCGGGTGGTCAACCTGGCC 61
Qy      2137 GlnHisSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSer 2156
Db      62 CAGCACATCAGTGGGTATCACACAGACTACACCCGGCACACCCACACAGCAGCTCAGC 121
Qy      2157 AlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeu 2176
Db      122 GCACCCCTGCGCGCCCTCTACTCTCTCTCTGGGCCAGTCTGCCCGTCTGGACCTC 181
Qy      2177 ArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGly 2196
Db      182 CGCGGCCACACAGTACCTTACCTCCGCGCCCGGACCATGTGTGCCCGCGCGCTGGC 241
Qy      2197 SerProHisSerGluGlyGlyValArgSerProGluProAsnLysThrSerValLeuGly 2216
Db      242 TCCCCCCCACAGGAGGGGGGAGAGGTCTCCAGAGCCAAACAGAGCTCGTCTTGGGT 301
Qy      2217 GlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHis 2236
Db      302 GGTGGTGGGACGGTATTGACCTGTGTCCACCGGAGGGCATGACGGAGCCAGGGCAC 361
Qy      2237 SerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArg 2256
Db      362 TCCCGAGTGTGTGTATCCCGCTGTGTACCGGATGGGGAACAGACAGCCAGCAGG 421
Qy      2257 MetGlySerLysSerProGlyAsnThrSerGlnProProAlaPheSerLysLeuThr 2276
Db      422 ATGGGCTCCAGTCTCCAGGCAACACCCAGCAGCGCGGCGAGCTTCTTCAGCAAGCTGACC 481
Qy      2277 GluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThr 2296
Db      482 GAGAGCACTCGGCATGTTCAAGTCCAGAGCAAGCAAGATCAACAGAGCTGAACACC 541
Qy      2297 HisAsnArgAsnGluProGluTyrAsnLysSerGlnProGlyThrGluIlePheAsnMet 2316
Db      542 CACAACCGGAATAGCCTGAATACAAATATACAGCAGCCTGGGAGCGAGATCTTCAATATG 601
Qy      2317 ProAlaIleThrGlyThrGly 2323
Db      602 CCGGCCATCACCGGACAGGT 622

```

```

RESULT 70
BF530324/c
LOCUS 602071630F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4214551
DEFINITION 5', mRNA sequence.
ACCESSION BF530324
VERSION BF530324.1 GI:11617698
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM9788 row: 1 column: 08
High quality sequence stop: 657.
Location/Qualifiers
1..657
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4214551"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sali; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Alignment Scores:
Pred. No.: 1.18e-36 Length: 657
Score: 1105.00 Matches: 210
Percent Similarity: 99.53% Conservative: 0
Best Local Similarity: 99.53% Mismatches: 1
Query Match: 8.36% Indels: 0
DB: 10 Gaps: 0
US-09-522-753-5 (1-2517) x BF530324 (1-657)

Qy      1447 LysGluGlySerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThr 1466
Db      657 AAGGAGGGCTCCATCAGCAGGGCACCCCGCTCAAGTACACACCCGCGGTCCACCACT 598
Qy      1467 GlySerLysLysHisAspValArgSerLeuIleGlySerProGlyArgThrPheProPro 1486
Db      597 GGCTCCAAAAGACAGACGACGCTCCCTCATCGGACAGCCCGCGCGGAGGATCCACCC 538
Qy      1487 ValHisProLeuAspValMetAlaAspAlaLeuGluArgAlaCysTyrGluGlu 1506
Db      537 GTGCACCGCTGGATGTGATGGCGACGCCCGGACCTGGAACGCTGCTGTACAGAGAG 478
Qy      1507 SerLeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAla 1526
Db      477 AGCTTGAAGACCGCGGACGAGCGCCAGCAGCTCGGGGGGCTCCATTGGCGCGGGCGCC 418
Qy      1527 ProValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHis 1546
Db      417 CCGGTCAATTGTGCTGAGCTGGGACCGCGGAGAGCCCTTAACCTATGAGGACCA 358
Qy      1547 GlyAlaProPheAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThr 1566
Db      357 GGGCACCTTTGCGGCGCACCTCCACAGAGTTTCGCCCGTGACACGCGGAGGCCACG 298
Qy      1567 ProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThr 1586
Db      297 CCGCGCTGCAGGAGGAGCGCTTTCTGTCCAGCAAGGCATCCCGAGGACCCGAAAGCTGACG 238
Qy      1587 SerThrProArgGluIleAlaLysSerProHisSerThrValProGluHisHisProHis 1606
Db      237 TCGAGCGCTGTGAGATCGCCAGTCCCGCACAGCAGCCGTGCCCGAGGACCCACAC 178
Qy      1607 ProIleSerProTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSer 1626
Db      177 CCCATCTCGCCCTATGACACCTCTCTCGGGCGCTGAGTGGCGTGGAGCTGTATCGCAGC 118
Qy      1627 HisIleProLeuAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAla 1646
Db      117 CACATCCCGCTGGCTTCGACCCCGACCTCCATATCCCGCGGCATCCCTCTGGAGCGAGCC 58
Qy      1647 AlaAlaTyrTyrLeuProArgHisLeuAlaPro 1657
Db      57 GCTGCTACTACTACCTGCCCCGACACCTGGCCCCC 25

RESULT 71
BG485444/c
LOCUS BG485444
DEFINITION 602505328F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4618783 5',
mRNA sequence.
ACCESSION BG485444

```

```

VERSION      BG485444.1  GI:13417723
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 902)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-remail.nih.gov
             Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCMI377 row: g column: 08
             High quality sequence stop: 629.
FEATURES     Location/Qualifiers
             1..902
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:4618783"
             /lab_host="DH10B (T1 phage-resistant)"
             /clone_lib="NIH MGC 77"
             /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
             SfiI (ggcccggccgccc); Site 2: SfiI (ggccattatggcc); 5' and
             3' adaptors were used in cloning as follows: 5' adaptor
             sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
             5'-ATTCTAGAGGCGGAGCGCGGACATG-dt(30)BN-3' (where B = A,
             C, or G and N = A, C, G, or T). Average insert size 1.9
             kb (range 0.5-4.0 kb). 12/15 colonies contained inserts.
             by PCR. This library was enriched for full-length clones
             and was constructed by Clontech Laboratories (Palo Alto,
             CA). Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      1.75e-36      Length:      902
Score:          1104.50      Matches:    252
Percent Similarity: 79.94%      Conservative: 7
Best Local Similarity: 77.78%      Mismatches: 31
Query Match:    8.36%      Indels:     36
DB:             12          Gaps:        5

US-09-522-753-5 (1-2517) x BG485444 (1-902)
Qy 2103 LeuProHis-----LeuArgProLeuProGluSerGlnProSerSerProLeu 2119
Db 890 CTACCAACTGGCGCGCTGAGACACGCCCT-----TGTCACCGCGG 846
Qy 2120 LeuGlnThrAlaProGlyValGlyGlyHisGlnArgValThrLeuAlaGlnHisIle 2139
Db 845 CTCCAGGCGCGCCCGCAGGTTCAAGGTTTCAACACCGGGTGTCCACTGCGCCAGCAAT 786
Qy 2140 SerGluValIle-ThrGlnAspTyrThrArgHis-HisProGlnGlnLeuSerAla---P 2158
Db 785 CATGAGGTGATCAACACAGGACTACAGCGCGGCACCAACCCACAGCAGCTCAGCGCAACC 726
Qy 2158 roLeuProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeu---AspLeuA 2177
Db 725 CTGCCCCCGCCCCCTGACGCGCGGTGCGCCAGCGCGCTGCGCGGTGCGCGGACCTGCGCG 666
Qy 2177 rGArgProProSerAsp-LeuTyr-LeuProProProAsp-HisGlyAlaPro-AlaArg 2195
Db 665 CCGCACAAACCCAGTGACCGGTACGGTCCCGCGCGGACCAATGGTCCCGGGCGCGT 606
Qy 2196 GlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThrSerValLeu 2215

```

```

Db 605 GGCTCCCCCAGCAGCAAGG-GGCAAGAGGTCTCCAGAGCCCAACAGACGTCGT--CTT 549
Qy 2216 GlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThrGluProGly 2235
Db 548 GGTGCTGTTGAGGACGGTATTGAACCTGTGTCCACCGGAGGGCATGACGGAGCCAGGG 489
Qy 2236 HisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSer 2255
Db 488 CACTCCCGGAGTGTGTGTACCGCTGTGTACCGGATGGGGAACAGACGAGGCCAGC 429
Qy 2256 ArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPheSerLysLeu 2275
Db 428 AGGATGGGCTCCAGTCTCCAGGCAACACACGAGCCGCGCCAGCTTCTTCAGCAAGCTG 369
Qy 2276 ThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsn 2295
Db 368 ACCGAGAGCAACTCCGCCATGTGTCAGTCCCAAGAGCAAGAGATCAACAAGAGCTGAAC 309
Qy 2296 ThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsn 2315
Db 308 ACCCACACACCGGAATGAGCCTGAATACAGC----- 279
Qy 2316 MetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHis 2335
Db 278 -----CTTATGACCTATTAGAGCCAGCGGTGAGGAACAT 243
Qy 2336 AlaSerThrAsnMetGlyLeuGluAlaIleAlaArgLysAlaLeuMetGlyLysTyrAsp 2355
Db 242 GCCAGCACCAACATGGGGCTGGAGGCCATATTAGAAAGGCACCTCATGGTAAATATGAC 183
Qy 2356 GlnTrpGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAla 2375
Db 182 CAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGAATGCCAGTGCC 123
Qy 2376 SerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThr 2395
Db 122 AGCTGCGCGCTGTGTATGCCCATACCGCTCTGACGAGCGAGTGACACACTCACC 63
Qy 2396 SerProGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLys 2415
Db 62 TCGCCAGGTGGCGGGGAGGCCAGGTCTCTGCGAGACCCAGCAGCGCGAAAGCCAG 3

RESULT 72
CB216442
LOCUS     NISC_nq05b08.y1_NICHDS_HS_Ut2_Homo_sapiens_cDNA_clone_IMAGE:5938118
DEFINITION 5', mRNA sequence.
ACCESSION CB216442
VERSION    CB216442.1  GI:28264634
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 625)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          cDNA Library Preparation:
          NCI-CGAP Arrayed by: The I.M.A.G.E. Consortium/LLNL
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC)
          Clone distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          info@image.llnl.gov
          Plate: LLCMI3166 row: C column: 15
          Seq primer: M13RP1 reverse primer (ABI).
FEATURES  Location/Qualifiers
             1..625
             /organism="Homo sapiens"

```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5938118"
/sex="female"
/tissue_type="normal endometrium, mid-secretory phase,
cycle day 23"
/lab_host="DH10B (TI-resistant)"
/clone_lib="NICHD HS Ut-2"
/note="Organ: uterus; Vector: pCMV-SPORT6.1.cdb (ResGen,
Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
Cloned unidirectionally from microquantity amounts of mRNA
from normal endometrial tissue (mid-secretory phase, cycle
day 23). Average insert size 1.6 kb. Library constructed
by ResGen (Invitrogen Corporation)."
```

ORIGIN

Alignment Scores:

Pred. No.:	1.88e-36	Length:	625
Score:	1099.50	Matches:	208
Percent Similarity:	99.52%	Conservative:	0
Best Local Similarity:	99.52%	Mismatches:	0
Query Match:	8.32%	Indels:	1
DB:	14	Gaps:	1

US-09-522-753-5 (1-2517) x CB216442 (1-625)

Qy 2163 LeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProProSerAsp 2182

Db 1 CTCTACTCTTCCCTGGGCGCAGCTGCCCCCTCTGGACCTCCGCCGCCACCCAGGTGAC 60

Qy 2183 LeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGly 2202

Db 61 CTCCTACTCTCCCGCCCCCGGACGTCGTCGCCCGCGCTGCTCCCCCAGCGAGGG 120

Qy 2203 GlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyLe 2222

Db 121 GGCAAGAGGTCTCCAGAGCCAAACAGACGTCGCTTGGTGGTGGTGGAGCGGTATT 180

Qy 2223 GluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyr 2242

Db 181 GAACCTGTGTCCCAACCGGAGGATGACGAGCGGGCACTCCCGAGTGTGTGTAC 240

Qy 2243 ProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerPro 2262

Db 241 CGCTGCTGTACCGGATGGGAGACAGACGAGGCC---AGGATGGGCTCCAGTCTCCA 297

Qy 2263 GlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMet 2282

Db 298 GGCAACACACGACGCGCCAGCCTTCTCAGCAAGCTGACCGAGAGCAACTCCGCCCATG 357

Qy 2283 VallysSerLysLysGlnGluLeuAsnLysLysLeuAsnThrHisAsnArgAsnGluPro 2302

Db 358 GTCAAGTCCAAAGAGCAAGATCAACAGAGGCTGAACACCCACACCGGAATGAGCCT 417

Qy 2303 GluTyrAsnLysSerGlnProGlyThrGluLeuPheAsnMetProAlaIleThrGlyThr 2322

Db 418 GAATACATATTCAGCAGCCTGGGAGCGAGATCTTCATATATGCCCGCATCACCGAACA 477

Qy 2323 GlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeu 2342

Db 478 GGCCTTATGACCTATAGAACCGCGCGTGCAGGAACATGCGCAGCACCAACATGGGGCTG 537

Qy 2343 GluAlaIleAlaArgLysAlaLeuMetGlyLysTyrAspGlnTyrGluGluSerProPro 2362

Db 538 GAGGCGCAATATAGAAAGCGCACTCATGGTAAATATGACCAGTGGGAAGAGTCCCCCGCG 597

Qy 2363 LeuSerAlaAsnAlaPheAsnProLeu 2371

Db 598 CTCAGCGCCATGCTTTTAACCCCTCTG 624

RESULT 73

BX732681

LOCUS

825 bp mRNA linear EST 18-NOV-2003

DEFINITION BX732681 XGC-tadpole silurana tropicalis cDNA clone TTPA071k07 5', mRNA sequence.

ACCESSION BX732681

VERSION BX732681.1

KEYWORDS GI:38405422

SOURCE EST.

ORGANISM Silurana tropicalis (western clawed frog)

Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Silurana.

1 (bases 1 to 825)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J. Sanger Xenopus tropicalis EST project 2001 (11_2003)

Unpublished (2003)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

TROPICALIS SEQUENCE ID: TTPA071k07.plkSP6

Sequencing primer: SP6

This sequence is from a Xenopus Gene Collection (XGC) library constructed by Nigel Garrett.

cDNA was oligo dt primed from sug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.

Vector: pCS107; Site_1: EcoRI; Site_2: NotI

Host: Escherichia coli DH10B.

Location/Qualifiers

1..825

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TTPA071k07"

/dev_stage="tadpole (stage 35-40)"

/lab_host="E. coli DH10B"

/clone_lib="XGC-tadpole"

/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dt primed from sug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Alignment Scores:

Pred. No.:	2.81e-36	Length:	825
Score:	1098.50	Matches:	216
Percent Similarity:	85.82%	Conservative:	20
Best Local Similarity:	78.55%	Mismatches:	38
Query Match:	8.31%	Indels:	1
DB:	13	Gaps:	1

US-09-522-753-5 (1-2517) x BX732681 (1-825)

Qy 138 AspArgSerLeuThrGlyLysLeuGluProValSerProProSerProProHisThrAsp 157

Db 3 GACAGGGGTTTGGCCAGCAACATGAGACTACTTTCGTCACTCTTTCTGGGCAACCTGGA 62

Qy 158 ProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuGlnAsnMetAsp 177

Db 63 GAGGAACAGGAGGCTCTCCATCTTAAATCTTCTTAAAGAGAACTTATCAAGTATGGAC 122

Qy 178 ArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysGln 197

Db 123 CGCTTTCACCGCAGATTGCTTAAAGTTGACAGCAATCTTAAAGTTGAAGAAAAACAG 182

Qy 198 GlnGlnLeuGluGluAlaAlaLysProGluProGluLysProValSerProPro 217

Db 183 CAACAGCTTGAAGAAGAGCAGCAAAACCTCCAGCCAGAGAGCCCGCTCTCCCTCCA 242

Qy 218 ProfileGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLys 237

Db 243 CCTGTGGAACAGAAGCATCGCAGTATAGTCCAAATTTATTTATGATGAAACCGGAAAAA 302

QY 238 AlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyr 257
 Db 303 CGAGAGGAGCACAATAGATCTTAGAGGCTTAGGTCCAAAAGTAGAGCTGCTCTTTAC 362

QY 258 AsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArg 277
 Db 363 AACCCAGCCTCAGACAAAGATTATCATGAAAACATCAGACCAATCAGGTGATGAGG 422

QY 278 LysLysLeuIleLeuTyrPhelysArgArgAsnHisAlaArgLysGlnTrpLysGlnLys 297
 Db 423 AAAAACAATACTCTCTTTAAGAGAGAATAATCATGTAGAAAACCTAAGGGACAGAAC 482

QY 298 PheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGlu 317
 Db 483 ATTTGTCAGCGCTATGATCAGCTGATGAGGCTTGGGAGAAAGATTGACAGGATAGAA 542

QY 318 AsnAsnProArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPhe 337
 Db 543 AATAATCTCGGAGAAAGCAAGAGAGAGCAAAACCCGAGAAATATATGAGAAGCAATT 602

QY 338 ProGluIleArgLysGlnArgGluLeuGlnArgMetGlnSerArgValGlyGlnArg 357
 Db 603 CCTGAATAAGAACAGACAGAGAGCAGCAGGAGCGTTTCCAG--AGGTTGGCAGAGA 659

QY 358 GlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleLeuAsp 377
 Db 660 GGAGCTGGCTCTCAGCAACCATGACCCAGGAGTGAACATGAATAATCTGAAATATTGAT 719

QY 378 GlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProPro 397
 Db 720 GGTCTTTCTGAACAGGAGAAATAATGAAAGCAGATGCGCCCACTCTCAGTCATTCCACCC 779

QY 398 MetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMet 412
 Db 780 ATGATGTTTGATCGGAGCANAGGAGGTGTAATTCATTAATG 824

RESULT 74
 BE542336
 LOCUS 601067079F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3453499 5',
 DEFINITION mRNA sequence.
 ACCESSION BE542336
 VERSION BE542336.1 GI:9770981
 KEYWORDS EST.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM8436 row: m column: 20
 High quality sequence stop: 701.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3453499"
 /cell_line="MGC36"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_10"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;

ORIGIN

Alignment Scores:
 Score: 7.14e-36 Length: 922
 Pred. No.: 1090.00 Matches: 236
 Percent Similarity: 70.72% Conservative: 8
 Best Local Similarity: 68.41% Mismatches: 25
 Query Match: 8.25% Indels: 77
 DB: 10 Gaps: 3

US-09-522-753-5 (1-2517) x BE542336 (1-922)

QY 80 GlnGluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGlu 99
 Db 1 CAGGAGCTCCACCTGCGGCCAGAGTCCACATCATCTACCTGCCCCGAGCTGGGAAAGTCAGAG 60

QY 100 MetGluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArg 119
 Db 61 ATGGAGTTCAATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCTGACCCCTGCTGCGA 120

QY 120 ProSerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArg 139
 Db 121 CCGTCACCCCTGCTGGCCAGCGGCGAGCTCGGGATCTGAAGACCTCACCAGGACCGT 180

QY 140 SerLeuThrGlyLysLeuGluProValSerProSerProProHisThrAspProGlu 159
 Db 181 AGCTGACGGCAAGCTGGAACCGGTGTCCTCCCGAGGCCCCAGACACTGACCCCTGAG 240

QY 160 LeuGluLeuValProProArgLeuSerLysGluGluLeuGlnAsnMetAspArgVal 179
 Db 241 CTGGAGCTGTGTCGCCACCGCTGTCCAGGAGGAGCTGATCCAGAACATATGACCGCGTG 300

QY 180 AspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysGlnGlnGln 199
 Db 301 GACCGAGAGATCACCTGTTAGAGCAGCAGATCTTAAGCTGAAGAGAGAGCAGCAACAG 360

QY 200 LeuGluGluAlaAlaLysProGluLysProGluLysProValSerProProProIle 219
 Db 361 CTGGAGGAGGAGGCTGCCAAGCCCGGAGCTGAGAAAGCCCGTGTACCCCGGCCATC 420

QY 220 GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysAlaGlu 239
 Db 421 GAGTCGAAGCACCAGCCCTGTTGACAGATCATCTACGACGAGAACCGG----- 468

QY 240 AlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGln 259
 Db 469 -----CCGCTGTACAACCCAG 483

QY 260 ProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLys 279
 Db 484 GCCTCCGACACCCCGCAGTATCATGAGAACATCAAAATAAAGAGGCGCATCGGAAAGAG 543

QY 280 LeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCys 299
 Db 544 CTATCTTTGACTTCAAGAGGAGGAGTACCGCTCGGAAACATATGGAGCAGAGATTCTTGC 603

QY 300 GlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsn 319
 Db 604 CAGCGCTATGACCAAGCTCATGGAGCGCTGGGAGAGAACCGTGGAGCGCATCGAGAACAC 663

QY 320 ProArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu 339
 Db 664 CCCGG-CCGCGGGCCAGGAGAGC----- 686

QY 340 IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySer 359
 Db 686 ----- 686

QY 360 GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu 379

Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5 kb. Library prepared by Life
 Technologies."

Db 686 ----- 686
 Qy 380 SerGluGlnGluAenLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu 399
 Db 687 ----- 687
 Qy 400 Tyr---AspAlaAspGlnArgIle-LysPheIleAsnMet-AsnGlyLeuMetAlaA 418
 Db 741 GTACGCAACGCTGACCGGCTACAGTGTCCATGTACGAGGCAATGGCCA 800
 Qy 418 spProMetLys 421
 Db 801 AGCCCGGAAG 811

RESULT 75
 LOCUS BX869832/c
 DEFINITION BX869832 AGENAE Rainbow trout multi-tissues-normalized (tcbk)
 ONCORHYNCHUS mykiss CDNA clone tcbk0022c.i.19 5prim, mRNA sequence.

ACCESSION BX869832.1 GI:39993419
 VERSION BX869832
 KEYWORDS ONCORHYNCHUS mykiss (rainbow trout)
 SOURCE ONCORHYNCHUS mykiss
 ORGANISM ONCORHYNCHUS mykiss
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 1 (bases 1 to 682)
 GOVOROUN, M., GUIGUEN, Y. and LE GAC, F.
 Construction and primary characterization of normalized cdna
 libraries in rainbow trout, Oncorhynchus mykiss
 Unpublished (2003)
 JOURNAL Contact: Guiguen Y
 COMMENT INRA - SCRIBE
 Campus de Beaulieu, RENNES cedex, 35042, France
 Tel: 02.23.48.50.09
 Fax: 02.23.48.50.20
 Email: Yam.Guiguen@beaulieu.rennes.inra.fr
 Sequence cleaned of vector, adaptor and repetitions. Contact us
 at signenaupport@jouy.inra.fr to obtain the chromatogram of this
 sequence.
 Plate: 0022 row: i column: 19
 Seq primer: M13R.

FEATURES
 source
 1..682
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="tcbk0022c.i.19"
 /tissue_type="multi-tissues"
 /dev_stage="from embryos to adults"
 /lab_host="DH10B"
 /clone_lib="AGENAE Rainbow trout multi-tissues-normalized (tcbk)"
 /note="Vector: pTT3D-pac; Rainbow trout multi-tissues-normalized + 2 subtractions; Clone distribution : AGENAE Resource centre. Francois PIUMI, Francois.PIUMI@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du genome (UREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73"

ORIGIN
 Alignment Scores:
 Pred. No.: 6,82e-36 Length: 682
 Score: 1087.00 Matches: 206
 Percent Similarity: 96.88% Conservative: 11
 Best Local Similarity: 91.96% Mismatches: 7
 Query Match: 8,23% Indels: 0
 DB: 13 Gaps: 0

US-09-522-753-5 (1-2517) x BX869832 (1-682)

Qy 241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
 Db 679 GCSCACAGGATTCCTGGAGGACTGGGACCCCGAGTAGAACTGCTCTGTACACACGCA 620
 Qy 261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
 Db 619 TCTGACACCAAGCAGTACCATGACATCAAAATAAAGCAGCGCATGAGGAAGAACTC 560
 Qy 281 ILeuTyrPheLysArgArgAsnHisAlaAsqLysGlnTyrLysGlnLysPheCysGln 300
 Db 559 ATTCTTTACTTCAAGAGGAGAAACCATGCTCGTAAGCAGTGGGACAGAAAGATTCTGCCAG 500
 Qy 301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
 Db 499 CGCTATGACCATCTGATGGAGGCTGGGAGAGAGAGTGGAGCGCATTTAGAGAACACCCC 440
 Qy 321 ArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
 Db 439 CGACGGCGGCTAAGGAGAGCAAGGTCCGAGAGTACTACGAGAAACAGATTCCCTGAGATC 380
 Qy 341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgLysSerGly 360
 Db 379 CGCAAGCAGAGAGAGCTGCAGGAGCGCATCGAGCGCGGGCGGCGGCGGAGGA 320
 Qy 361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluLysLysLeuAspGlyLeuSer 380
 Db 319 CTGGCCTCATCTGCAGCCCGCAGTGAACACAGAGTCTCTGAGATCGTTGACGGCATATCA 260
 Qy 381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
 Db 259 GAGCAAGAGAGAACACCGAGAGCAAAATCGCCAGCTGCGCGTCATCCCTCCCATGTTGTTT 200
 Qy 401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
 Db 199 GACCCGAGCAGCAGCGCATCAAGTTCATCAATGACCGCGCTGATGGACGCCCATG 140
 Qy 421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
 Db 139 AAGGTGTACAAGACCGCGCAGGTTCATGAACATGTGGAGCGAAACAGAGGAAGGACACCTTC 80
 Qy 441 ArgGluLysPheMetGlnHisProLysAsnGlyLeuIleAlaSerPheLeuGluArg 460
 Db 79 CGCGAGAAGTTCATCCAGCACCCCAAAACATTTGGCCTGATTGCTTCTTCTCTCGAGAGA 20
 Qy 461 LysThrValAla 464
 Db 19 AAGACGGTGGCC 8

RESULT 76
 LOCUS BI081252
 DEFINITION 602879279F1 NCI_CGAP_Mam2 Mus musculus cdna clone IMAGE:5011003 5', mRNA sequence.
 ACCESSION BI081252
 VERSION BI081252.1 GI:14499582
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 790)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Pred. No.: 1.3e-35 Length: 729
 Score: 1081.00 Matches: 208
 Percent Similarity: 92.80% Conservative: 11
 Best Local Similarity: 88.14% Mismatches: 16
 Query Match: 8.18% Indels: 2
 DB: 14 Gaps: 0

US-09-522-753-5 (1-2517) x CA325207 (1-729)

Qy 1402 LeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAla 1421
 Db 23 CTGGGTCCCTTGAAGCTGAAGCCCACTCACAGAGGTGTGGTAGC-ACTGTGAAGAGGCG 81
 Qy 1422 GlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuProLeu 1441
 Db 82 GGCCGCTTATCCATGAGATCCGAGAGAGAGTGGCCGACACTGAGTACCTCCCTG 141
 Qy 1442 AlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAspThr 1461
 Db 142 GCACCAAGGCTCTGAAGAGGGTTCCATCACCCAGGCGACCCCACTCAAGTACGACTCT 201
 Qy 1462 GlyAlaSerThrThrGlySerLysHisAspValArgSerLeuIleGlySerProGly 1481
 Db 202 GGGGCACTCCACTGGCACCAGAAACACACAGCTGGCTCCATCATCGGAGCCCGGC 261
 Qy 1482 ArgThrPheProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGluArg 1501
 Db 262 CGGCTTTCCTCCGCTGACCCGCTGACATATAGCTGACGCGCGGCACTGGAGCGT 321
 Qy 1502 AlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGlySer 1521
 Db 322 GCCTGCTATGAAGAGTCTGAAGAGCGGTTCAGGACCCAGTGTGTGAGGCGGCTCC 381
 Qy 1522 IleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerProLeu 1541
 Db 382 ATCACAGTGGGGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
 Qy 1542 ThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProValThr 1561
 Db 442 ACTTACGAAGACACAGGGGACCTTCCACAGTCCCTGCTGCTGCTGCTGCTGCTGCT 501
 Qy 1562 MetArgGluProThrProArgLeuGlnGlySerLeuSerSerSerLysAlaSerGln 1581
 Db 502 ACGAGGAGCCACCGCCAGCTTCAGAGAGCGCTTCCTATCCAGCAAGCGCTCCAG 561
 Qy 1582 AspArgLysLeuThrSerThrProArg-GluIleAlaLysSerProHisSerThrValPr 1601
 Db 562 GACCGAAGCTCACATCTACACCCCGGAGAGTCCCAAGTCCCAAGTCCCAAGTCCCA 621
 Qy 1601 cGluHisProHisProIleSerProTyrGluHisLeuArgGlyValSerGlyVa 1621
 Db 622 CGAGCACCACTCCACCTCCCTATGAGCATTGCTCCGGGCGTGACTGGTGT 681
 Qy 1621 LasLeuTyrArgSerHisIleProLeuAlaPheAspProThrSer 1636
 Db 682 GGACCTGTACCTGTGTCATCCCATTTGGCTTTGACCCCACTCC 727

RESULT 78

BU475754
 LOCUS 603471578F1 CSEQRBN22 809 bp mRNA linear EST 30-NOV-2002
 DEFINITION 603471578F1 CSEQRBN22 Gallus gallus cDNA clone ChEST348120 5', mRNA sequence.

ACCESSION BU475754

VERSION BU475754.1 GI:25969331

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 809)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,

Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

12445392

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. .809

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Layer and broiler"

/db_xref="taxon:9031"

/clone="CHEST348120"

/sex="Male and female"

/tissue type="Chondrocytes isolated from growth plate cartilage"

/dev_stage="adult"

/lab_host="DH10B"

/clone lib="CSEQRBN22"

/notes="Vector: pBluescript II KS(+); Site 1: EcoRI;

Site 2: NotI; This normalized library was constructed from

1 million independent clones. cDNA synthesis was initiated

using an oligo(dT) primer, using methylated C in the first

strand synthesis reaction. Following this first strand

reaction, double-stranded cDNA was blunted, ligated to

NotI adapters, digested with EcoRI, size-selected, and

cloned into the NotI and EcoRI compatible sites of a

custom modified MCS of the pBluescript (KS+) vector. The

library was normalized in 2 rounds using conditions

adapted from Soares et al., PNAS (1994) 91: 9228-9232 and

Bonaldi et al., Genome Research 6 (1996): 791, except that

a significantly longer reannealing hybridization was

used."

ORIGIN

Alignment Scores:

Pred. No.: 2.84e-35 Length: 809
 Score: 1074.00 Matches: 230
 Percent Similarity: 85.87% Conservative: 13
 Best Local Similarity: 81.27% Mismatches: 17
 Query Match: 8.13% Indels: 25
 DB: 13 Gaps: 4

US-09-522-753-5 (1-2517) x BU475754 (1-809)

Qy 334 GluLysGlnPheProGluIleArgLysGlnArgGluLeuGlnArgMetGln----- 351
 Db 4 GAGAAGCAATTCCTGAGATCCGGAAGCAGCGGAACTGCAAGAGCGCATGAGGTTA 63
 Qy 352 -----SerArgValGlyGlnArgGlySer---GlyLeuSerMetSerAlaAArgSer 368
 Db 64 GAGGAGAAGCGGTAGACACAGAGGGGCGGAGGGCTCTCCATGTCGGCTGCCGCGAGT 123
 Qy 369 GluHisGluValSerGluIleIleAspGlyLeuSerGluGlnGlnAsnLeuGluLysGln 388
 Db 124 GAGCACGAGGTGTCGGAGATCATGATGGGCTCTCAGAGCAAGAGAACCTGGAGAGCAG 183
 Qy 389 MetArgGlnLeuAlaValIleProProMetLeuTyrAspAlaAspGlnArgIleLys 408
 Db 184 ATGCCCGAGCTCGCCGTCATCCCATCCATGCTATGCTGACCGAGCGCATATAA 243
 Qy 409 PheIleAsnMetAsnGlyLeuMetAlaAspProMetLysValTyrLysAspArgGlnVal 428
 Db 244 TTCAATCAATCAACGCGCTCATGGACGCCCATGAGGTCTACAGGACCGCGCGGTG 303
 Qy 429 MetAsnMetTyrSerGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisPro 448


```

DEFINITION 601286040F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3607929 5',
mRNA sequence.
ACCESSION BE391091
VERSION BE391091.1 GI:9336456
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 688)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaops-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
CLONE DISTRIBUTION: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM259 row: p column: 10
High quality sequence start: 8
High quality sequence stop: 684.
FEATURES
Location/Qualifiers
1..688
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3607929"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_44"
/notes="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCACAGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

Alignment Scores: 3,32e-35 Length: 688

Pred. No.: 1070.50 Matches: 222

Score: 98.24% Conservative: 1

Percent Similarity: 97.80% Mismatches: 4

Best Local Similarity: 8.10% Indels: 4

Query Match: 10 Gaps: 0

DB:

US-09-522-753-5 (1-2517) x BE391091 (1-688)

QY 82 LeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMetGlu 101
Db 6 CTCACCTCGCGCCAGAGTCCCACTCATCTGCCGAGCTGGGAGCTCAGATGGAG 65

QY 102 PheLeuSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgProSer 121
Db 66 TTCATTAAAGCAAGCGCCCTCGGTAGAGTGTCTGCCCTGACCCCTCTCTCGACCGTCA 125

QY 122 ProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeu 141
Db 126 CCCCTGCTGGCCACGGGCGAGCTCGGGAGTCTGAGACCTTCCCAAGAGCCGTAGCTG 185

QY 142 ThrGlyLysLeuGluProValSerProSerProHisThrAspProGluLeuGlu 161
Db 186 ACGGCAAGCTGAACCGGTGTCTCCCCAGCCCGCCGACCTGACCTGAGCTGGAG 245

QY 162 LeuValProProArgLeuSerLysGluLeuLeuLeuGlnAsnMetAspArgValAspArg 181
Db 246 CTGTGTCGCGCCACGGCTGTCTCAAGAGGAGGCTGATCCAGAACATGGACCGGTGGACCGA 305

QY 182 GluileThrMetValGluGlnGlnIleSerLysLeuLysLysGlnGlnGlnLeuGlu 201
Db 306 GAGATCACCATGCTAGAGCAGCAGATCTCTAAGCTCAGAGAAGCAGCAAGCTGGAG 365

QY 202 GluGluAlaAlaLysProGluProGluLysProValSerProProProlleGluSer 221
Db 366 GAGGAGGCTGCCAAGCGCCGAGCGCCGCTGTCCACCGCCGCTATCGAGTCG 425

QY 222 LysHisArgSerLeuValGlnIleLeuTyrAspGluAsnArgLysLysAlaGluAlaAla 241
Db 426 AAGCACCAGCAGCTGTGTGAGATCATCTACGAGCAGAACCGGAAGAGCTGAAGCTGCA 485

QY 242 HisArgileLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrHisGlnProSer 261
Db 486 CATCGGATTCTGGAAGGCTTGGGGCCCGAGGTGCGGTGTGTACACCAAGCCCTCC 545

QY 262 AspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLeuLeu 281
Db 546 GACACCCGCGAGTATCATGAGAACATCAATA-AACGAGCGCATGCGGAAGAGCTAATC 604

QY 282 LeuTyrPhelysArgAsnHisAlaArgLysGlnTyrLysGlnLysPheCysGlnArg 301
Db 605 TTGTACTTCAAGAGGAGGAT-CACGCTCGGAACA-TGGGAGCAGAGATC-TGCCAGCGC 661

QY 302 TyrAspGlnLeuMetGluAla 308
Db 662 TATGACCAGCTCATGGAGGCT 682

RESULT 81
BF058781/c 664 bp mRNA linear EST 16-OCT-2000
LOCUS 7K34907.x1 NCI CGAP Ov18 Homo sapiens cDNA clone IMAGE:3477325 3'
DEFINITION similar to TR:00613 000613 T3 RECEPTOR-ASSOCIATING COFACTOR-1. [1]
; mRNA sequence.

ACCSSION BF058781 GI:10812677
VERSION BF058781.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 664)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaops-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
CLONE DISTRIBUTION: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 467.
Location/Qualifiers
1..664
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3477325"
/tissue_type="fibrothecoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI-CGAP Ov18"
/notes="Organ: ovary; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGCATTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Alignment Scores:
 Pred. No.: 3,35e-35 Length: 664
 Score: 1070.00 Matches: 201
 Percent Similarity: 98.07% Conservativity: 2
 Best Local Similarity: 97.10% Mismatches: 4
 Query Match: 8,10% Indels: 0
 DB: 10 Gaps: 0

US-09-522-753-5 (1-2517) x BF058781 (1-664)

Qy 1902 SerThrSerProValArgProAlaAlaThrPheProProAlaThrHisCysProLeu 1921
 Db 622 TCACCTTCATACCTGTCGCAGGCTGCCATTCCTCCCTTGCACCATGTCACATG 563
 Qy 1922 GlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLysGlu 1941
 Db 562 ACAGGACCATCATGGGTCTACCTTACCCTCATGGAGCCGCTTCTGCTGCCAAGGAG 503
 Qy 1942 AlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeuAla 1961
 Db 502 GCGCCCGGGTCGCCCGCCAGAGCGGCCGAGCAGACACCGGCCATGCTTCTCGCC 443
 Qy 1962 LysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGluPro 1981
 Db 442 AAGCCCCCAGCCCGCTCGGGCTGGAGCCGCTCTCCCTCCAGCAAGGGCTCGAGGCC 383
 Qy 1982 ArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsn 2001
 Db 382 CGGCCCATAGTGCCTCTCTCTGCGCCAGCCACCATCGCCGACCCCTCGAAGAAC 323
 Qy 2002 LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAspPro 2021
 Db 322 CTCGCACCTCACCGCCAGCCGCGGACCGCGCGCCACCTGCTCGGCTCGGACCG 263
 Qy 2022 HisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeu 2041
 Db 262 CACCGGGAAGACTCAAGTAACCCCTTTTCATCCAGGAACCTGGAACCTCGGTCTCTG 203
 Qy 2042 GlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSerSer 2061
 Db 202 GGTACACGCGCAGCAGCTACAGCCCGAGGGGTGGAGCCCGTCAGCCCTGTGAGCTCA 143
 Qy 2062 ProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHis 2081
 Db 142 CCCAGTCTGACCCACGACGAGGGGCTCCCAAGCACCTGGAAGAGCTCGACAAGAGCCAC 83
 Qy 2082 LeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAla 2101
 Db 82 CTGGAGGGGAGCTGCGGCCCAAGCAGCAGCCGCGCCGCTGGAAGCTTGGCGGGAGCGGCC 23
 Qy 2102 HisLeuProHisLeuArgPro 2108
 Db 22 CACCTCCACACCTGCGGCGG 2

RESULT 82
 BUI08220
 LOCUS BUI08220 1088 bp mRNA linear EST 25-NOV-2002
 DEFINITION 603112484F1 CSEQCHL12 Gallus gallus cDNA clone ChEST61922 5', mRNA
 sequence.
 ACCESSION BUI08220
 VERSION BUI08220.1 GI:25311097
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 1088)

AUTHORS

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 2235534
 12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1. 1088
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="ChEST61922"
 /dev_stage="36"
 /lab_host="DH10B"
 /clone_lib="CSEQCHL12"
 /note="Organ: heart; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
 [Stratagene] vector to accommodate cDNA produced with the
 T-trimmed protocol (Construction of uni-directionally
 cloned cDNA libraries from messenger RNA for improved 3'
 end DNA sequencing by Glenn Fu, et al. U.S. Patent #
 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
 Ligate in double stranded adaptor containing BglI and
 BamHI sites [5'ggccgcgtcagccgcatcggaaaaaag]
 [5'aattcttttttcggatccggggtgcagc]"

ORIGIN

Alignment Scores:
 Pred. No.: 1,18e-34 Length: 1088
 Score: 1062.50 Matches: 235
 Percent Similarity: 74.79% Conservativity: 35
 Best Local Similarity: 65.10% Mismatches: 64
 Query Match: 8.04% Indels: 29
 DB: 13 Gaps: 9
 US-09-522-753-5 (1-2517) x BUI08220 (1-1088)
 Qy 1117 ValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValPro 1136
 Db 1 TTGCTGGAGAGACCTGTGGCTCCATCTCCAGGGAATGCCCATCCAGCTCCACACCC 60
 Qy 1137 Tyr---SerGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuPro 1155
 Db 61 TACAGCTGTGAGCATGCAAAAGTTCCTGTCTGGCTCCATCCATGATGGGCTGCCACTGACA 120
 Qy 1156 MetAspProLysLysLeuAlaProPheSerGlyValLysGlnGlnLeuSerProArg 1175
 Db 121 ATGGATCCCAAGAACTGGTGCCTTTCTCGAGTAAGCAGGAACAACTCTCTCCCGAGA 180
 Qy 1176 GlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeu 1195
 Db 181 AGCCAGGCGCCAGCCAGCTGAAAGCTGTGTATGATGAGCGGCCCGGAGGGCTCGCTCTG 240
 Qy 1196 ArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyLysProSerThr 1215
 Db 241 CGGGGTAGCTCCCTCAGCTCTGCTCAGGAGGAGCATCACCAAGGGGACACCCATCC 300
 Qy 1216 ArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAla 1235
 Db 301 AGACCCCTCCAGAGTCGCCCATCAGTACCGAGGCTCCATCACACGCGCACCCCGCGCA 360
 Qy 1236 AspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeu 1255
 Db 361 GAAGTGTGTACAAAGGAACCATTTACCAGGATAATCCGAGAGATAATCCGAGAGAGAGCCG 420


```

BI735071      820 bp      mRNA      linear      EST 20-SEP-2001
LOCUS      603556210F1 NTH_MGC_94 Mus musculus cDNA clone IMAGE:5363282 5',
DEFINITION      mRNA sequence.
ACCESSION      BI735071
VERSION      BI735071.1 GI:15712084
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      1 (bases 1 to 820)
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L2AW11924 row: d column: 03
High quality sequence stop: 725.
Location/Qualifiers
1..820
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5363282"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
Note: "Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      7,74e-34      Length:      820
Score:      1039.50      Matches:      213
Percent Similarity:      87.35%      Conservative:      8
Best Local Similarity:      84.15%      Mismatches:      25
Query Match:      7.87%      Indels:      9
DB:      12      Gaps:      3

US-09-522-753-5 (1-2517) x BI735071 (1-820)

Qy      1929      TyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgPro 1948
Db      1      TACCCTACCTCATGAGCCCGCTCTGTATCCCAAGAGAGACCTCTCGGGTGCCTCGGCC 60
Qy      1949      GluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProAlaArgSerGly 1968
Db      61      GAGCGGCCCCGTGTGAGCTGGCCATGCTTCTCCACCAACCCCGGCCGG----- 114
Qy      1969      LeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProVal 1988
Db      115      ---GAGCCCGCTCTCTACCCAGCAGAGCTCCGAGCCCGCATCCCTAGCACCCCGCAGC 171
Qy      1989      SerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHialaSer 2008
Db      172      TCCAGCCACACAGCCATCGCGCCAGCCACCAAGAAACCTTGCACCCACCATGCTCAGT 231
Qy      2009      ProPheProAlaProAlaSerAlaSerAlaSerProHisArgGluLysThrGlnSer 2028
Db      232      CCGGACCCCGCGGGG---CCCACTTCGGCCTCAGATCTGCACCGGAGAAAGACTCAAGT 288
Qy      2029      LysProPheSerIleGlnGluLeuArgSerLeuGlyTyrHis---GlySerSer 2047

```

```

Db      289      AAACCCCTTTTCATCCAGGAATTGGAACCTCGTTCTCTGGGTATACACAGTGGAGCTGC 348
Qy      2048      TyrSerProGluGlyValGluProValSerProValSerProSerLeuThrHisAap 2067
Db      349      TACAGCCCCGATGGGTGGAGCCCATCAGCCGCGGTAGCTCCCCAGGCTGACCCACGAC 408
Qy      2068      LysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGluLeuArg 2087
Db      409      AAGGGGCTCTCCAAACCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAAGGGAGCTGCG 468
Qy      2088      ProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArg 2107
Db      469      CACAAGCAGCCAGGCCCATGAAGCTCAGCGCGAGGCTGCCCATCTCCCATCTCGCGG 528
Qy      2108      ProLeuProGluSerGlnProSerSerProLeuLeuGlnThrAlaProGlyValLys 2127
Db      529      COACTGCCCGAGAGCCAGCCCTCATCTAGCCACATCTCTCCAGACTGCCCGAGCATCAA 588
Qy      2128      GlyHisGlnArgValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyr 2147
Db      589      GGTCAACAGAGGGTGGTCACTTGGCTCAGCACATCAGTGGAGTCACTTACGAGGACTAC 648
Qy      2148      ThrArgHisHisProGlnGlnLeuSer--AlaProLeuProAlaProLeuTyrSerPheP 2167
Db      649      ACGGCGCCACGA-CGCGCAGCAGCTCAAGTGGCCCTTCCCGCCCTCTCTACTCTCTTC 707
Qy      2167      roGlyAlaSerCysProValLeuAspLeuArgArg 2178
Db      708      CCGGAGCAGCTGG-CCTGTCCGGGATCTTCGGCAG 741

RESULT 85
LOCUS      AW438580
DEFINITION      x982e10.x1 NCI CGAP Ut2 Homo sapiens cDNA clone IMAGE:2776170 3'
similar to TR:000613 O00613 T3 RECEPTOR-ASSOCIATING COFACTOR-1. [1]
, mRNA sequence.
ACCESSION      AW438580
VERSION      AW438580.1 GI:6973886
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 611)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moshaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco
High quality sequence stop: 420.
Location/Qualifiers
1..611
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2776170"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Ut2"

```

FEATURES
source

/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.85 kb. Life Technologies catalog #: 11539-012"

ORIGIN

Alignment Scores:

Pred. No.: 6,46e-34 Length: 611
Score: 1038.00 Matches: 194
Percent Similarity: 96.06% Conservatively: 1
Best Local Similarity: 95.57% Mismatches: 8
Query Match: 7.85% Indels: 0
DB: 10 Gaps: 0

US-09-522-753-5 (1-2517) x AW439580 (1-611)

Qy 2118 ProLeuLeuInThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGln 2137
Db 3 CCGCTGCTCCAGACCGCCAGGGGTCAAGGTCAACAGGTCAACAGCGGGTGGTCAACCTGGCCAG 62
Qy 2138 HisLeuSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAla 2157
Db 63 CACATCAGTGGGTCAATCACAGAGCTACACCCGGCCAGCCACCCAGAGAGCTCAGCGCA 122
Qy 2158 ProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArg 2177
Db 123 CCCCCTGCGCGCCCTCTACTCTCTCTGGGGCCAGCTGCCCGTCTGGACCTCCGC 182
Qy 2178 ArgProProSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySer 2197
Db 183 CGCCACCCAGTACCTTACCTCCCGCCCGGACCATGTGTGCCCGCCGTGGCTCC 242
Qy 2198 ProHisSerGluGlyGlyValArgSerProGluProProLeuValThrSerValLeuGlyGly 2217
Db 243 CCCCACAGCGAAGGGGGGAGAGGCTCCAGAGCCAAAGAGCTCGTCTTGGTGGT 302
Qy 2218 GlyGluAspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSer 2237
Db 303 GGTGAGGACGGTATTGAACCTGTGTCACCGGAGGCGCATGACGGAGCCAGGCACTCC 362
Qy 2238 ArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnInThrGluProSerArgMet 2257
Db 363 CGAGTGTGTGTATCCCGCTGTGTACCGGATGGGAAACAGAGAGCCAGCAGGATG 422
Qy 2258 GlySerLysSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGlu 2277
Db 423 GCCTCAAGTCTCCAGGCAACACAGCCAGCCGCGCCAGCTTCTTCAAGCACTGACCGAG 482
Qy 2278 SerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHis 2297
Db 483 AAGCACTCCGTATGTCAGTGCAGAGCAAGCAAGAGATCAACAGAGCTGAACACTCAC 542
Qy 2298 AsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetPro 2317
Db 543 AACCGGAATAGCCTGAATACAATATTCGCCAGCCCTGGAGCGGGAATCTTAAATATGCC 602
Qy 2318 AlaIleThr 2320
Db 603 GGCATCACC 611

RESULT 86

BG391632

LOCUS

DEFINITION 602417893F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4537438 5',

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 624)

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-x@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lni.gov>
Plate: LHAM10462 row: i column: 23
High quality sequence stop: 623.

FEATURES

source

1..624
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4537438"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 6,61e-34 Length: 624
Score: 1038.00 Matches: 206
Percent Similarity: 99.04% Conservatively: 0
Best Local Similarity: 99.04% Mismatches: 2
Query Match: 7.85% Indels: 2
DB: 12 Gaps: 0

US-09-522-753-5 (1-2517) x BG391632 (1-624)

Qy 1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
Db 3 AGCGCCATCATACATCCCGCGCTCCATCACCAGCGCACCGCAGCTGACGTCCTGTACAG 62
Qy 1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
Db 63 GGCACCATCATCCAGGATCATCGCGGAGGACAGCCCGAGTCGTTGGACCCGCGCGGAG 122
Qy 1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
Db 123 GACAGCTGCCAAGGCCACGTCATCTAGCAAGGCAAGAAGGGCCACGCTTGTCTCTAT 182
Qy 1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerGlyPro 1300
Db 183 GAGGTGGCATGTCTGTGACCCAGTGTCTCAAGAGGAGCGGAGGAGGAGGAGGAGGAGG 242
Qy 1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
Db 243 CCCCATGAGCGCGCCGCCACAGCGCACCTATGACATGATGGAGGCGCGGTGGGCGAGA 302
Qy 1321 AlaIleSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
Db 303 GCCATCTCTCAGCAGCATCGAAGGTCTCATGGCGGTGTCATCCCGCGCGGAGGAGGAGG 362
Qy 1341 SerProHisLeuLysGluGlnHisIleArgGlySerIleThrGlnGlyLeuPro 1360
Db 363 AGCCCCCACCACCTCAAGAGGAGCAGCACCATTCGCGGGTCCATCACACAGGAGATCCCT 422
Qy 1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuArg 1380
Db 423 CGGTCTACGTGGAGGACAGGAGGAGTACTCTGCTCGGAGGAGGAGGAGGAGGAGGAGG 482
Qy 1381 GluGlyThrProProProProProProProProProProProProProProProPro 1400

Db 483 GAGGGACGCTCCCGCCCGCCCTCACGGGACCTGACGGGCGCTTACAGGACGCGAG 542
 1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
 Db 543 GCCTTGGGCGCCCTGAGCTGAGCGCGC-CATGAGGGGCGCTGGTGGCCACCGTGAAGGAG 601
 1421 AlaGlyArgSerIleHisGluLeu 1428
 Db 602 GCGGGCGCTCC-ATCCATGAGATC 624

RESULT 87
 CB518985 668 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-GH0-gee-f-04-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
 DEFINITION IMAGE:6838373 5', mRNA sequence.
 ACCESSION CB518985
 VERSION
 KEYWORDS EST.
 SOURCE CB518985.1 GI:29352340
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 668)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousef.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..668
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6838373"
 /tissue_type="whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_GH0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CCACTGAAT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

FEATURES

source
 Alignment Scores:
 Pred. No.: 7,126-34 Length: 668
 Score: 1038.00 Matches: 199
 Percent Similarity: 93.30% Conservative: 10
 Best Local Similarity: 88.84% Mismatches: 13
 Query Match: 7.85% Indels: 2
 DB: 14 Gaps: 0

ORIGIN

Alignment Scores:
 Pred. No.: 7,126-34 Length: 668
 Score: 1038.00 Matches: 199
 Percent Similarity: 93.30% Conservative: 10
 Best Local Similarity: 88.84% Mismatches: 13
 Query Match: 7.85% Indels: 2
 DB: 14 Gaps: 0

US-09-522-753-5 (1-2517) x CB518985 (1-668)
 QY 1418 ValLysGluAlaGlyArgSerIleHisGluLeuProArgGluGluLeuArgHisThrPro 1437
 Db 1 GTGAAGGAGGGCGGGCGCTCTATCATGAGATCCCGAGAGAGGAGCTGGCCGCGCACCN 60
 1438 -GluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeu 1457
 Db 61 TGAGCTACCCCTGGCACCCACGCGCTCTCGAAGGAGGGTTCCATCACCCAGGCGACCCACT 120
 QY 1457 uLysTyRAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuL 1477
 Db 121 CAAGTACGACTCTGGGGCACCTCCACTGGGCACCAAGAAACACGACGTCGCTCCATCAT 180
 QY 1477 eGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAla 1497
 Db 181 CGGAGAGCCCGCGCGCTTTCCCTGGCCCTGACCCGCTGGACATATGGCTGACGCCG 240
 QY 1497 gAlaLeuGluArgAlaCysTyRArgGluSerLeuLysSerArgProGlyThrAlaSerSe 1517
 Db 241 GGCAGTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCGGTCAGGACCGCAGTG 300
 QY 1517 rSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProAr 1537
 Db 301 TGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGCTGCTGAACTGGGCGACGCCAG 360
 QY 1537 gGlnSerProLeuThrTyRArgHisGlyAlaProPheAlaGlyHisLeuProArgG 1557
 Db 361 GCAAGGCCACTGACTTACGAAGAGCCACGGGGACCTTCCAGTACCTGCCACGTGG 420
 QY 1557 ySerProValThrMetArgGluProThrProArgLeuGlnGlySerLeuSerSerSe 1577
 Db 421 CTCCCTGTGACACGAGGAGCCACGCCACGCTTCAGGAGGACGACCTCTCTATCCAG 480
 QY 1577 rLysAlaSerGlnAspArgLysLeuThrSerThrProArg-GluLeuAlaLysSerProH 1597
 Db 481 CAAGCGCTCCCGAGGACCGAAGCTGACATCTACACCCCGGAGATGCCAAGTCCCCAC 540
 QY 1597 iSerThrValProGluHisProHisProHisProHisProHisProHisProHisProHis 1617
 Db 541 ACAGCACTGTCCCGGAGCACCACCTCACCCATCTCCCTCCCTATGAGACATTCCTCCG 600
 QY 1617 lyValSerGlyValAspLeuTyRArgSerHisIleProLeuAlaPheAspProThrSerI 1637
 Db 601 GCGTGACTGGTGTGGACCTGTACCGTGTGTCATCCATCCCATTCGCTTTCACCCCTCCA 660
 QY 1637 leProArg 1639
 Db 661 TACCCCGA 668
 RESULT 88
 CD244013 925 bp mRNA linear EST 22-MAY-2003
 LOCUS AGENCOURT 14069335 NIH MGC 180 Homo sapiens cDNA clone
 DEFINITION IMAGE:30377707 5', mRNA sequence.
 ACCESSION CD244013
 VERSION CD244013.1 GI:31004477
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 925)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM442 row: f column: 20

High quality sequence start: 21

High quality sequence stop: 560.

Location/Qualifiers

FEATURES

source

```
1..925
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30377707"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:
Pred. No.: 1,12e-33 Length: 925
Score: 1037.00 Matches: 239
Percent Similarity: 64.34% Conservative: 10
Best Local Similarity: 61.76% Mismatches: 31
Query Match: 7.85% Indels: 108
DB: 14 Gaps: 9

US-09-522-753-5 (1-2517) x CD244013 (1-925)

```
QY 2146 AspTyrThrArgHisGlyProGlnLeuSerAlaPro----LeuProAlaProLeuTyr 2164
DB 18 GAATCTTGGGATGCATCGACCTGTCCAAGTGCACACACCTGCTGTGTCGCC 77
QY 2165 SerPheProGlyAla-----SerCysPro 2172
DB 78 CCGACACAGGACCCCGACGCCGATGAGCCGCTGTGCTACCTCCCGAGTGGCCC 137
QY 2173 ValLeuAspLeuArgProSerAspLeuTyrLeuProProAspHisGlyAla 2192
DB 138 GTCTTGACCTCCGCGCCACCCAGTACCTCTACCTCCGCGCCCGGAGCATTGTGCC 197
QY 2193 ProAlaArgGlySerProHisSerGluGlyLysArgSerProGluProAsnLysThr 2212
DB 198 CCGGCGGTGGTCCCGCCACAGCGAAGGGGCAAGAGTCTCCAGAGCCAAACAGAG 257
QY 2213 SerValLeuGlyGlyGluAspGlyIleGluProValSerProGluGlyMetThr 2232
DB 258 TCGTCTTGGTGGTGGTGGAGACGGTATTGAACCTGTGTCGCCACCGAGGGCATGAG 317
QY 2233 GluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGlnThr 2252
DB 318 GAGCCAGGCGACCTCCCGAGTGTGTGTACCGGCTGTGTACCGGATGGGGAACAGAG 377
QY 2253 GluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProAlaPhePhe 2272
DB 378 GAGCCAGCAGGATGGGCTCCAGTCTCCAGGCAACCCAGGAGCCAGCCACCTTCTTC 437
QY 2273 SerLysLeuThrGluSerAsnSerAlaMetValLysSerLysGlnGluLeuLys 2292
DB 438 AGCAAGCTGACCGAGAGCACTCCGCGATGGTCAAGTCCCAAGAGCAAGAGATCAACAG 497
QY 2293 LysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnLysSerGlnProGlyThrGlu 2312
DB 498 AAGCTGAACACCCACCAACCGAATGAGCCTGATACATATCAGCCAGCTGGGCGGAG 557
QY 2313 IlePheAsnMetProAlaIleThrGlyLeuMetThrTyrArgSerGlnAlaVal 2332
DB 558 ATCTCAATATGCGCCGATCACCAGAAC----- 587
QY 2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352
```

```
DB 587 ----- 587
QY 2353 LysTyrAspGlnTyrGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372
DB 587 ----- 587
QY 2373 AlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHis 2392
DB 587 ----- 587
QY 2393 ThrLeuThrSerProGlyGlyLysAlaLysValSerGlyArgProSerSerArg 2412
DB 588 -----GGTGGCGCGGAGGAGGCTCTCTGGCAGACCAAGC-AGCCGA 631
QY 2413 LysAlaLysSerProAlaPro-GlyLeuAlaSerGlyArgPro-ProSerValSerS 2432
DB 632 AAGCCCAAGTCCCGCGCCCGGGGCTGTCATCTGGGACCGGCCACCTCTGTCTCT 691
QY 2432 erValHisSerGlu-GlyAspCysAsnArgArg-ThrProLeu-ThrAsnArgValTyrG 2451
DB 692 CAGTGCACCTCGGAAGGAGACTGCAACCGCGGAGCGCGCTCAACAAACCGGTGTGGG 751
QY 2451 luAspArg---ProSerSer-AlaGlySerThrProPhePro---TyrAsnProLeuIle 2468
DB 752 AAAGAAAGAGGCGCTCTGTTCCCGAGGTTCCCGCCATTCCTTACAAACCCCTGAATC 811
QY 2469 MetArgLeuGlnAlaGlyValMet-----AlaSerProProProProGly---LeuPro 2485
DB 812 ATCGGGGTTCAGAGCGCGGTGTGTCAATGGCTTCCCGACCCCGCGGGGCGCTTCCC 871
QY 2486 AlaGly-----SerGlyProLeuAlaGlyProHisAlaTyrasp-GluGluProLy 2503
DB 872 CCCGGCGCAGCGGGGCGCCCTT-----CGCTGGGCGCCCGCCCAACAG 916
QY 2503 sProLeu 2505
DB 917 GCCCTTG 923

RESULT 89
LOCUS BG119261 1030 bp mRNA linear EST 30-JAN-2001
DEFINITION 602349134F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4443969 5',
mRNA sequence.
ACCESSION BG119261 GI:12612767
VERSION BG119261
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1030)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Prepared by: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10219 row: c column: 10
High quality sequence stop: 625.
Location/Qualifiers
1..1030
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4443969"
/tissue_type="adenocarcinoma, cell line"
```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 1-53e-33 Length: 1030
Score: 1035.00 Matches: 213
Percent Similarity: 86.38% Conservative: 9
Best Local Similarity: 82.8% Mismatches: 32
Query Match: 7.83% Indels: 5
DB: 12 Gaps: 2

US-09-522-753-5 (1-2517) x BGI19261 (1-1030)

Qy 2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGluAsp 2220
Db 1 GAAGGGGGCAGAGGTCTCCAGAGCCAAACAGACGTCGGTCTTGGTGGCGTGGAGC 60
Qy 2221 GlyIleGluProValSerProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
Db 61 GGTATTGAACCTGTCTCCACCGGAGGCGATGACGGAGCCAGG-CACTCCCGGAGTGCT 119
Qy 2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
Db 120 GGTATACCGCTCTGTACCGGATGGGAAACAGAGCCAGAGCCAGAGATGGCTCCAG 179
Qy 2261 SerProGlyAsnThrSertGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
Db 180 TCTCCAGGCAACACCAGCAGCGCCGAGCTTCTTCAGCAAGCTGACCGAGCAACTCC 239
Qy 2281 AlaMetValLysSerLysGlnGluLeuAsnLysLysLeuAsnThrHisAsnArgAsn 2300
Db 240 GCCATGGTCAAGTCCAAAGAGAGATCAACAGAGAGCTGAACACCCCAACACCGGAAT 299
Qy 2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
Db 300 GAGCTCGAATACAAATATCGCCAGCGCTGGAGCGGAGATCTTCAATATGCCGCCATCACC 359
Qy 2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsn-Me 2340
Db 360 GGAACAGGCTTATGACCTATAGAGCCAGCGGCTGCAGACATGCCAGACCAACATT 419
Qy 2340 tGlyLeuGluAlaIleArgLysAlaLeuMetGlyLysTyrAspGlnTyrGluGluSe 2360
Db 420 GGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCATGGGAAGATC 479
Qy 2360 rProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAl 2380
Db 480 -CCGCGCTCAGCGGCAATGCTTTTAAACCTCTGAATGCCAGTCCAGATCGCGTCT 538
Qy 2380 aMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGly 2400
Db 539 ATGCCCATTAAGCTGCTTGACGGACGGAGTGACCAACCAATCACTTGGCCAGTGGCGG 598
Qy 2400 tGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProG 2420
Db 599 CGGGAAGGCCAAGGTCTCTGGCAGACCAAGAGCG--AAAGCCCAAGTCCCGGCCCGGG 655
Qy 2420 yLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAs 2440
Db 656 CTTGGCATCTGGGGACCGGACACCTCTCTCTCTCATGCACTCGAG--GAGACTGC 712
Qy 2440 nArgArgThrProLeuThrAsnArgValTyrGluAspArgProSerSer 2456
Db 713 AGCGCTAGGCGGTACAAACCCCGTTTGGCAGACACAGGCCCTATTC 761

RESULT 90
AW701437

LOCUS AW701437 654 bp mRNA linear EST 18-APR-2000
DEFINITION uq6b05.y1 NCI CGAP Lu33 Mus musculus cDNA clone IMAGE:2937969 5'
similar to TR:O00613 O00613 T3 RECEPTOR-ASSOCIATING COFACTOR-1. [1]
; mRNA sequence.

ACCESSION AW701437
VERSION AW701437.1 GI:7585568
SOURCE EST.
KEYWORDS Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (Bases 1 to 654)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

FEATURES

source

MG1:1050389
Seq primer: -40RP from Gibco
High quality sequence stop: 470.
Location/Qualifiers
1..654
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:2937969"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI-CGAP_Lu33"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES

source

Alignment Scores:
Pred. No.: 1.07e-33 Length: 654
Score: 1033.50 Matches: 205
Percent Similarity: 78.49% Conservative: 3
Best Local Similarity: 77.36% Mismatches: 10
Query Match: 7.82% Indels: 47
DB: 10 Gaps: 2
US-09-522-753-5 (1-2517) x AW701437 (1-654)

ORIGIN

Qy 2188 ProbAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyLysArgSerPro 2207
Db 1 CCCGACCATGCTCCCCAGCCCGGGATCCCCCAACAGTGAAGGGGGCAAGGTCCTCCCA 60
Qy 2208 GluProAsnLysThrSerValLeuGlyGlyGluAspGlyIleGluProValSerPro 2227
Db 61 GAACCCAGCAAAACATCGTCTCTGGCAGCAGTGGAGTGCATTCAGCTGCTGTGCCCA 120
Qy 2228 ProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArg 2247
Db 121 CCAGAGGGGCATGACTGAGCCAGGACATGCTCGAGCGCTGTGTATCCCACTGCTGTATCGA 180

```

Qy 2248 AspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGln 2267
Db 181 GACGGGACAGCGGAGCCC--AGATGGCTCTAAGTCTCCAGGCAACACAGCCAG 237
Qy 2268 ProProAlaPheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLys 2287
Db 238 CCGCCAGCCTTCTTTCAGTAAGCTGACTGAGAGCAACTCCGCGATGGTGAAGTCGAAGAAG 297
Qy 2288 GlnGluLeuLeuLeuLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSer 2307
Db 298 CAGGAGATCAACAAGAAACTCAACCCACCAACCGGACAGCCAGATCAATATATTGGC 357
Qy 2308 GlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyr 2327
Db 358 CAGCTGGGACGGAAATCTTCAACATGCCGCCCATCACTGGAGCAGGCCATTATGACCTGT 417
Qy 2328 ArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleArg 2347
Db 418 AGAAGCCAGGCGGTGCAAGAAACAGCCAGCACCACCAATGCGGCTAGAGGCCATTATTAGA 477
Qy 2348 LysAlaLeuMetGlyLysTyrAspGlnTrpGluSerProProLeuSerAlaAsnAla 2367
Db 478 AAGCACTCATG----- 489
Qy 2368 PheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAsp 2387
Db 489 ----- 489
Qy 2388 GlyArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGly 2407
Db 490 -----GGTGGAGGTGGGAAAGGCAATGTCTCTGGC 519
Qy 2408 ArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgPro 2427
Db 520 AGACCTAGCCCGGAAAGCCAGGTCGCACAGCCAGCCCTAGCGTCTGAGACCGACCC 579
Qy 2428 ProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsn 2447
Db 580 CTTCTGTCTACTGATGACTCAGAGGGGACTGCATCCCGGACACCACTACCAAC 639
Qy 2448 ArgValTrpGluAsp 2452
Db 640 CGTGTGTGGAAGAC 654

RESULT 91
CA317171 725 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FW0-cbm-p-15-0-UI:r1 NIH_BMAP_FW0 Mus musculus cDNA clone
DEFINITION IMAGE:6811744 5', mRNA sequence.
ACCESSION CA317171
VERSION CA317171.1 GI:24535295
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 725)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-romail.nih.gov
Tissue procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.

```

FEATURES

```

Location/Qualifiers
1..725
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6811744"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FW0"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

```

ORIGIN

```

Alignment Scores: 1.38e-33 Length: 725
Pred. No.: 1032.00 Matches: 209
Score: 88.52% Conservative: 7
Percent Similarity: 85.66% Mismatches: 23
Best Local Similarity: 7.81% Indels: 7
Query Match: 14 Gaps: 3
DB:

```

US-09-522-753-5 (1-2517) x CA317171 (1-725)

```

Qy 1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
Db 7 TTGGTGGGACCCCTTGAGGGGCTTACCCTACCTCATGGAGCCCGTCTCTGTACCCAAG 66
Qy 1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
Db 67 GAGACCTCTCGGGTCGCCCGCCGCGCGCGCTGTGGACGCTGGCCATGCTTCTCTC 126
Qy 1961 AlalysProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
Db 127 ACCAAACCCCGGCCCGG-----GAGCCGCGCTCTCTACCCAGCAAGAGCTCCGAG 177
Qy 1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
Db 178 CCGCGATCCCTAGCACCCCGCCAGCTCCAGCCACACAGCCATCGCCGCGCCAGCAAG 237
Qy 2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
Db 238 AACCTTCACCCACCATGCGAGTCGCGAGCCGCGCGCG---CCACCTCGGCCCTCAGAT 294
Qy 2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
Db 295 CTGCACCCGAGAAAGACTCAAGTAACCCCTTTTCATCAGGAATTTGGAACTCCGTCT 354
Qy 2041 LeuGlyTyrHis---GlySerSerTyrSerProGluGlyValGluProValSerProVal 2059
Db 355 CTGGGTTACCACAGTGGAGCTGGCTACAGCCCGCGATGGGTGGAGCCCATCAGCCCGGTG 414
Qy 2060 SerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluLeuAspLys 2079
Db 415 AGCTCCCGCCAGCTGACCCAGCAAGGGGCTCTCCAAACCTCTGGAAAGCTAGAGAAG 474
Qy 2080 SerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlu 2099
Db 475 AGCCACTTTGAGAGGGAGCTGCCGCGACAGCCAGCCGCCCATGAGACTCAGCGCGGAG 534

```

QY 2100 AlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeu 2119
 |||||
 Db 535 GCTGCCCATCTCCACATCTGGCGCCACTGCGCCGAGAGCCN-CCCTCATCTAGCCCACTC 593
 |||||
 QY 2120 LeuGlnThrAlaProGlyValIysGlyHisGlnArgValValThrLeuAlaGlnHisIle 2139
 |||||
 Db 594 CTCACAGACTGCCCGCCAGGATCAAAAGGTCAACAGAGGGTGGTCACTTGGCTCAGCACATC 653
 |||||
 QY 2140 SerGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeu 2159
 |||||
 Db 654 AGTAGGTCATATACGACAGACTACACGCCGCCACACCCGCGACGACTCAGTGG-CCCCCTT 712
 |||||

QY 2160 ProAlaProLeu 2163
 |||||
 Db 713 CCGCGCCCTCTC 724

RESULT 92

BM771666

LOCUS

DEFINITION K-EST0055566 S7SNU719s1 Homo sapiens cDNA clone S7SNU719s1-21-D11
 5', mRNA sequence.

ACCESSION

BM771666

VERSION

BM771666.1

GI:19101281

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 614)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 21 row: D column: 11

High quality sequence stop: 614.

Location/Qualifiers

1..614

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S7SNU719s1-21-D11"

/sex="M"

/tissue_type="Stomach"

/cell_type="Epithelial"

/cell_line="SNU-719"

/lab_host="Top10P"

/clone_lib="S7SNU719s1"

/note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;

Site 2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including EcoR

I site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized from oligo dt-selected mRNA by

priming with dt-tailed vector. The dt-tailed vector was

adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of

EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The

obtained cDNA vectors were used for transformation of

competent cells E. coli Top10P by electroporation method.

The cDNA libraries constructed by this method are

full-length enriched cDNA library. After analyzing and

sequencing about 2,000 ~ 3,000 colonies in original cDNA

library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dt)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10P with electroporation method."

ORIGIN

Alignment Scores:
 Pred. No.: 1.39e-33 Length: 614
 Score: 1030.00 Matches: 203
 Percent Similarity: 72.50% Conservative: 0
 Best Local Similarity: 72.50% Mismatches: 1
 Query Match: 7.79% Indels: 76
 Db: 12 Gaps: 2

US-09-522-753-5 (1-2517) x BM771666 (1-614)

QY 2151 HisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSer 2170
 |||||
 Db 3 CACCCACAGCAGCTCAGCGCACCCCTGCCGCCCTCTACTCTCTCCCTGGGCGCAGC 62
 |||||
 QY 2171 CysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHis 2190
 |||||
 Db 63 TGCCCGCTCTCGACCTCCGCCGCCACCCAGTACCTCTACCTCCGCCCGCCGACCAT 122
 |||||
 QY 2191 GlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsn 2210
 |||||
 Db 123 GGTGCCCGCGCCCTGGCTCCCCACACGCGAGGGGCGACAGAGGTCTCCAGAGCGCAAC 182
 |||||
 QY 2211 LysThrSerValLeuGlyGlyGlyAspGlyLeuGluProValSerProProGluGly 2230
 |||||
 Db 183 AAGACGTCTGGTCTTGGGTGGCGGTGAGGACGGTATTGAACCTGTCTCCACCGAGGGC 242
 |||||
 QY 2231 MetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGlu 2250
 |||||
 Db 243 ATGACGGAGCCAGGGCCTCCCGGAGTGTCTGTACCCGCTGCTGTACCGGGATGGGAA 302
 |||||
 QY 2251 GlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAla 2270
 |||||
 Db 303 CAGACGGAGCCCC---AGGATGGGTCTCCAGTCTCCAGGCAACACAGAGCGCCAGGCC 359
 |||||
 QY 2271 PhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIle 2290
 |||||
 Db 360 TTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCCAAGCAAGAGATC 419
 |||||
 QY 2291 AsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGly 2310
 |||||
 Db 420 AACAGAAGCTGAACACCCACACCGGAATGAGCTGAATACAAATATATCGGCGAGCTGG 479
 |||||
 QY 2311 ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGln 2330
 |||||
 Db 480 ACGGAGATCTTCAATATATGCCCGCCATCACCCGGAACA----- 515
 |||||
 QY 2331 AlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeu 2350
 |||||
 Db 515 ----- 515
 |||||
 QY 2351 MetGlyLysTyrAspGlnTrpGluSerProProLeuSerAlaAsnAlaPheAsnPro 2370
 |||||
 Db 515 ----- 515
 |||||
 QY 2371 LeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSer 2390
 |||||
 Db 515 ----- 515
 |||||
 QY 2391 AspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArgProSer 2410

```
Db 516 -----GTTGGCGCGGGAAGCCAAAGGTCTCTGTGGCAGACCCAGC 554
Qy 2411 SerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerVal 2430
Db 555 AGCCGAAAGCAAGTCCCGCGCGCGGCTGGCATCTGGGGACCGCGCCACCTCTGTC 614

RESULT 93
BF606925
LOCUS BF606925 810 bp mRNA linear EST 01-APR-2001
DEFINITION MY2 000117 Mouse 9-day fetus cDNA library MPMGP559 Mus musculus
cDNA clone MPMGP559C0355 5', mRNA sequence.
ACCESSION BF606925
VERSION BF606925.1 GI:13503489
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 810)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B.G.,
Lehrach, H. and O'Brien, J.
TITLE Detection of a high number of novel genes in a 9-day mouse embryo
cDNA library normalised by oligonucleotide fingerprinting
JOURNAL Unpublished (2001)
COMMENT Contact: Hennig S
Laboraty 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr.63-73, D-14195 Berlin, Germany
Tel.: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
Randomly selected clones.
PCR Primers
FORWARD: 5'-GAGCTATTCAGAGAGTAGTGA-3'
BACKWARD: 5'-TAATACGACTCAGTATAGG-3'
Seq primer: 5'-ATTAGGTGACACTATAG-3'
High quality sequence stop: 810.

FEATURES
source
Location/Qualifiers
1..810
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MPMGP559C0355"
/tissue_type="whole embryo"
/dev_stages="embryonic 9-day"
/lab_host="E.coli DH10B Maxefficiency (Gibco cells)"
/clone_lib="Mouse 9-day fetus cDNA library MPMGP559"
/notes="Vector: pSVSPORT1; Site 1: NotI; Site 2: SalI;
Library preparation by oligo dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."

ORIGIN
Alignment Scores:
Pred. No.: 2,4e-33 Length: 810
Score: 1027.50 Matches: 215
Percent Similarity: 81.00% Conservative: 11
Best Local Similarity: 77.06% Mismatches: 36
Query Match: 7.78% Indels: 19
DB: 10 Gaps: 6

US-09-522-753-5 (1-2517) x BF606925 (1-810)
Qy 2088 ProLysGlnProGly---ProValLysLeuGlyGlyGluAlaHisLeuProHisLeu 2106
Db 18 CCGGAATTCGGGGTCGACCCACCGCT-----CCGCCACG 53
Qy 2107 ArgProLeuProGluSerGlnProSerSerProLeuLeuGlnThrAlaProGlyVal 2126
Db 54 CGTCCG-----CCCTCATCTAGCCACCTCCACAGCTGCCCGGAGGATC 98
Qy 2127 LysGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAsp 2146
```

```
Db 99 AAAGGTCAACAGAGGGTGGTCACTTGGCTCAGCACATCAGTGAAGGTCAATTACGACGAC 158
Qy 2147 TyrThrArgHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhe 2166
Db 159 TACACGGCGCCACCAACCCGCGAGCAGCTCAGTGGCCCCCTTCCCGCCCCCTCTCTACTCTTT 218
Qy 2167 ProGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuPro 2186
Db 219 CCGGAGCCAGCTGCGCTCTCTGGTCTTTCGCGCGCCACCCAGTGTCTCTACTCTCCA 278
Qy 2187 ProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSer 2206
Db 279 CCCCCGACCATGCGACCCCGCGGGATCCCGCCACAGTGAAGGGGGCAAAAGGTCC 338
Qy 2207 ProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSer 2226
Db 339 CCAGATCCCAACAAACATCGTCTCTGGGCGAGCAGCGAGGATGCCATGTAGACCTGTGTCC 398
Qy 2227 ProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyr 2246
Db 399 CCACGAGGGCGATGACTGAGCCAGGACATGCTCGGAGCGCTGGTATCCACTGCTGTAT 458
Qy 2247 ArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSer 2266
Db 459 CGAGACGGGGAACAGGGCGAGCCCC--AGGATGGGCTCTAAGTCTCCAGGCAACACGACG 515
Qy 2267 GlnProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLys 2286
Db 516 CAGCGCCGACGCTTCTTTCAGTAAGTCACTGAGAGCAACTCCCGCATGGTGAAGTCAAC 575
Qy 2287 LysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIle 2306
Db 576 AGCAGGAGATCAACATGAACATCATCCACACATCCGGTACGAGCCAGATACATATT 635
Qy 2307 SerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThr 2326
Db 636 GGCCAGCCTGGGACGGAAATCTTCAACATGCCCCCATCACTGGAGCAGGCGCTTATGTA- 694
Qy 2327 TyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuAlaIle 2346
Db 695 CTGTATTAAGCAGCGCGTGCATGAC---ACGACGACATCATCTGGCTAGAGGCATTATTA 751
Qy 2347 ArgLysAlaLeuMetGlyLysTyrAspGlnTyrGluSerProProLeuSerAla 2365
Db 752 TAAAGGACC---ATGGGTAATATGATCATCTGCGAAGAGCC-CGCGCGCTCGGCGCC 804

RESULT 94
AI523558/c
LOCUS AI523558 612 bp mRNA linear EST 13-APR-1999
DEFINITION th08C09.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2117680 3',
similar to TR:000613 000613 T3 RECEPTOR-ASSOCIATING COFACTOR-1. [1]
;contains MER22.b2 MSRI repetitive element ;, mRNA sequence.
ACCESSION AI523558
VERSION AI523558.1 GI:4437693
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 612)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
```


found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.lnl.gov/bbrp/image/image.html

Insert Length: 857 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 488.

FEATURES

Location/Qualifiers
1..612
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2117680"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_CLL1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTACCAATCTGAGTGGAGCGCCGACATGCTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 2,23e-33 Length: 612
Score: 1025.00 Matches: 198
Percent Similarity: 97.06% Conservatives: 0
Best Local Similarity: 97.06% Mismatches: 6
Query Match: 7.76% Indels: 1
DB: 9 Gaps: 0

US-09-522-753-5 (1-2517) x AF523558 (1-612)

Qy 1905 SerProValArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThr 1924
Db 612 TCACCCCGTTCGCGCGTGCACATTCACACN-GCCACCCATGCCACTGCGCGGCACC 554
Qy 1925 LeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArg 1944
Db 553 CTCGATGGGGTCTACCCCTACCTCATGAGCGCCGCTTGTGCGCCAGAGGCCCGCGTG 494
Qy 1945 ValAlaArgProGluArgProAlaAlaThrGlyHisAlaPheLeuAlaLysProPro 1964
Db 493 GTGCGCGCGCAGAGCGCGCGCGAGCAGACCGGCGCATGCTTCTCGCCAGGCCCA 434
Qy 1965 AlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeu 1984
Db 433 GCGCGCTCGGCGCTGGAGCGCGCGCTCTCCCGCAGCAAGGGTTCGAGCGCGCGCCCTA 374
Qy 1985 ValProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaPro 2004
Db 373 GTGCGCTCTGTCTCTGGCGAGCGCCACCATCGCGCGCGCCGCGCGCGCGCGCGCT 314
Qy 2005 HisHisAlaSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGlu 2024
Db 313 CACCAAGCGGAA 254
Qy 2025 LysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHis 2044
Db 253 AAGACTCAAGTAAACCTTTTCCATCCAGGAACCTGGAACCTCGCTCTCTGGGTTATCCAC 194
Qy 2045 GlySerSerTyrSerProGluGlyValGluProValSerProValSerSerProSerLeu 2064
Db 193 GCGACAGCTACAGCTG 134
Qy 2065 ThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGly 2084
Db 133 ACCACAGCAAGGGGCTCCCCAAGCAGCTGGAAGAGCTCGAAGAGCCACTGGAGGGG 74
Qy 2085 GluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaHisLeuPro 2104

Db 73 GAGCTGGCGGCCAAGCAGCAGGCTCCGTGAAGCTTGGCGGGAGGCGGCCACCTCCCA 14
Qy 2105 HisLeuArgPro 2108
Db 13 CACCTGGCGCGC 2
RESULT 95
BF339827
LOCUS
DEFINITION 602038834F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4186688
5', mRNA sequence.
ACCESSION BF339827
VERSION BF339827.1 GI:11286286
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 898)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Straubeberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9508 row: c column: 09
High quality sequence stop: 765.
FEATURES
Location/Qualifiers
1..898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4186688"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Brn64"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:
Pred. No.: 5.23e-33 Length: 898
Score: 1020.50 Matches: 214
Percent Similarity: 82.06% Conservatives: 33
Best Local Similarity: 71.10% Mismatches: 41
Query Match: 7.72% Indels: 15
DB: 10 Gaps: 5

US-09-522-753-5 (1-2517) x BF339827 (1-898)

Qy 194 LysLysLysGlnGlnGlnLeuGluGluAlaLysProGluProGluLysPro 213
Db 3 AAAAAGAAACAAACACAGCTTGAAGAAGAGCAGCTAAACCTCTGAGCCTGAGAAGCCC 62
Qy 214 ValSerProProLysGluSerLysHisArgSerLeuValGlnIleIleTyrAspGlu 233
Db 63 GTGTCCCTCTCTCTGTGGAGCAGAACACCGCAGTATTGTCCAAATTTATTATGATGAG 122
Qy 234 AsnArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGlu 253
Db 123 AATCGGAAAAAGCAGAGAAGCTCATTAATTTTGAAGCTCTTGCGCCCAAAAGTTGAA 182
Qy 254 LeuProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsn 273
Db 183 CTGCCACTGTATACCCAGCCATCAGATCAACCAAGGTGTACCATGAGAACATCAAGCAAAAC 242

Fax: 314 296 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
ReSourceZentrumPrümarbeitBank, Berlin, Germany (web address:

High quality sequence stop: 497.

FEATURES

source

```
/organism="Danio rerio"
```

/mol type="mRNA"

```
/db_xref="taxon:7955"
```

```
/clone="IMAGE:4967317"
```

```
/sex="mixed male and fe
```

```
/tissue_type="brain"
```

```
/dev_stage="adult"
```

```
. /lab_host="E. coli DH10
```

```
/clone_lib="zebrafish a
```

```
/note="Vector: pZIPLOX;
```

Original library was closed

excision of the cDNA in

pZ1PLOX plasmids. Inset

Library."

ORIGIN

Alignment Scores:

Assignment Scores:	
Pred. No.:	4.77e-33
Length:	637

REGA: NO.:	1703
Scores:	1017.50
MATCHES:	195

Percent Similarity:	Conservative:
96.24%	10

Best Local Similarity:	Mismatches:
91.55%	7

Query Match:	7.70%	Indels:	1
--------------	-------	---------	---

DB:	12	Gaps:	1
-----	----	-------	---

US-09-522-753-5 (1-2517) x BI428444 (1-637)

		260	ProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLys	270
Qy				
		280	LeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCys	290
Qy				
		300	GlnArgTyrAspGlnLeuMetGluAlaLeuGlnLysLysValGluArgIleGluAsnAsn	310
Qy				
		320	ProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu	330
Qy				
		340	IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySer	350
Qy				
		360	GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu	370
Qy				
		380	SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu	390
Qy				
		400	TyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro	410
Qy				
		420	MetLysValTyrLysAspArgGlnValMetAsnMetTyrSerGluGlnGluLysGluThr	430
Qy				
		440	ProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLys	450
Qy				
		460	LeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCys	470
Qy				
		480	GlnArgTyrAspGlnLeuMetGluAlaLeuGlnLysLysValGluArgIleGluAsnAsn	490
Qy				
		500	ProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu	510
Qy				
		520	IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySer	530
Qy				
		540	GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu	550
Qy				
		560	SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu	570
Qy				
		580	TyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro	590
Qy				
		600	MetLysValTyrLysAspArgGlnValMetAsnMetTyrSerGluGlnGluLysGluThr	610
Qy				

274	Qy	GlnAlaMetArgLysLysLeuIleuTyrPhe-LysArgArgAsnHisIslaArgLysG1	4291
243	Db	CAGGTGATCAGAGAAATCTCATTTATTTTACAAGAGAGAAATCATCTCAAGAAACA	302
293	Qy	nTrpLysGlnLysPheCysGlnATgTyrAspGlnLeuMetGluAlaLeuGluLysLysVa	313
303	Db	AGGGHACAAAAAATCTGCCAGCGTTATGATCAGCTCATGGAGGCATGGGAGAAAAAGT	362
313	Qy	lGluArgIleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTy	333
363	Db	GGACACAATAGAAAAATAATCTCTCGAGGAAAGCTAAAGAAAGCAAAACAAGGCAATCTA	422
333	Qy	rgLulLysGlnPheProGluIleArgLysGlnArgGluLeuGlnLuarGmetGlnSerAr	353
423	Db	TGAAACAGCAGTTTCAGAAAAATTCGAAAAACAAGAGAACACGACAGAAAGATTTCAG--CG	479
353	Qy	gValGlyGlnArgGlySerGlyLeuSerMetSerAlaalaArgSerGluHisGluValse	373
480	Db	AGTTGGGCACAGGGAGCTGGTCTTTTCAGCCACCAT-GCTAGAGTGAGCATGATGATTTTC	538
373	Qy	rgLulLeileAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAl	393
539	Db	TGAAATATTGATGGGCTCTCTGAGCAGGAGAAATAATCAGAAAACAAAATGCGCGAGCTCTC	598
393	Qy	aValIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAs	413
599	Db	TGTGATTCACCTATGATGTGTGATGCAGAACAGAGCAGTCCAGGTTCAATTAACTAGAA	658
413	Qy	nGlyLeuMetAlaAspProMetLysValTyrLys-AspArgGlnValMetAsnMetTrps	433
659	Db	TGGGCTTATGGAGGACCTATGAAAGTCTTTAACAGATAGGCAGTTTATGATGT-TGGA	717
433	Qy	erGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyL	453
718	Db	CTGACACAGACAGGACAATCTTTCAGGACAACAGTTAT--CAGGCATCCAAAAAATCTTGGAC	774
453	Qy	eulLeAlaSerPheLeuGluArgLysThrValAlaGluCysValLeu--TyrTyrTyrL	472
775	Db	TATGCTCTACCTGG-----GAAGAGAGTGTCCGATGGGTTCGTTTACAA	819
472	Qy	euthrLysLysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArgGly	491
820	Db	TTTAAACCAAGAAATAG-----ATTTAAGCTCGTCAGGAGATATCGAACCGCGGG	869

RESULT 96	BT428444	637 bp	linear	EST 26-JUL-2002
LOCUS	f8-8407.41	zebrafish adult	Tr:Q9WU43	Q9WU43
DEFINITION	IMAGE:4967317 5' similar to TR:Q9WU43 Q9WU43 SILENCING MEDIATOR OF RETINOIC ACID AND THYROID HORMONE RECEPTOR BETA. [1] ;, mRNA sequence.			

VERSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BI428444	BI428444	EST	Danio rerio (zebrafish)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
BI428444.1	BI428444.1	GI:15205676	Danio rerio (zebrafish)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS

1. F. (pages 1 to 637)

Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, J., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE	WashU Zebrafish EST Project 1998
JOURNAL	Unpublished (1998)
COMMENT	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800

479	ATGAAGTCTTACAAAGACAGACAGGTTCATGAACATGTGGAGCGAACAGGAGAGACACC	Db
440	PheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheIeuGlu	Qy
539	TTCCGTTGAGAAGTTTATCCAGACACCCAAAAACCTTGCCTGTATAGGCTCTTTTCTGGAG	Db

Qy 460 ArgLysThrValAlaGluCysValLeuTyrTyrTyrLeu 472
 599 AGAAGACGGTGGCTGAGTGTCTCTTCTATACCTG 637
 Db

RESULT	97
BU143568	
LOCUS	BU143568
DEFINITION	6038229009F1 CSOCHL26 Gallus gallus cDNA clone CHEST221b4 5', mRNA
LINEAR	linear
EST	EST 25-NOV-2002

sequence.
 BU143568
 ACCESSION
 BU143568.1
 VERSION
 GI:25360169
 KEYWORDS
 EST

SOURCE
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 760)

REFERENCE
AUTHORS
Boardman, P.B., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

TITLE
JOURNAL
MEDLINE
22335534
PUBMED
12445392

CONTACT	<p>Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk. Location/Qualifiers 1. .760</p>
FEATURES	source

```

organism="Oreochromis mossambicus"
mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST221b4"
/sex="Female"
/dev stage="adult"
/lab host="DH10B"
/clone_lib="CSEOCHL26"
notes="Organ: heart; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BglI and BamHI sites [5'gccgcgctgcagcccgatccgaataaaag] [5'aattcttttctggactcggagctgcagc]"

```

[illegible]

Db 62 AAAGTCCACCGGGTGTACCTCCGCGCCAGCACATCAGTGAGTGTATCACACGAGAC 121

Qy 2147 TyrThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhe 2166

Db 122 TACACCCCGGACACCCACGAGCTCAGCGCACCCCTCGCGCCCTCTACTCTCTTT 181

Qy 2167 ProGlyAlaSerCysProValLeuAspLeuArgProProSerAspLeuTyrLeuPro 2186

Db 182 CTTGGGGCGAGCTGGCCGCTCTGACCTCCGCGCCACCCAGCTGACCTCTACCTCCCG 241

Qy 2187 ProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyValysArgSer 2206

Db 242 CCCCCGACCATGGTCCCGCCGCTGGCTCCCGCCAGCAAGGGGCAAGAGGCT 301

Qy 2207 ProGluProAsnLeuThrSerValLeuGlyGlyGluAspGlyLeuProValSer 2226

Db 302 CCAGAGGCCAAACACAGACGCTGCTTGGTGGTGGTGGAGCGGTATTGAACCTGTGTC 361

Qy 2227 ProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyProLeuLeuTyr 2246

Db 362 CCACCGGAGGCGATCAGCGAGCCAGGCTCCCGAGTGTGTGTATCCCGTGTGTATC 421

Qy 2247 ArgAspGlyGluGlnThrClnProSerArgMetGlySerLysSerProGlyAsnThrSer 2266

Db 422 CGGGATGGGGAACACAGCGAGCCCATCAGGATGGGCTCCAGTCTCCAGGCAACCCAGC 481

Qy 2267 GlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLys 2286

Db 482 CAGCTCGAGCTTTTTCAGCGAGTGCAGCGAGCACTCCGCTCATGTCTAGTCCAG 541

Qy 2287 LysGlnGluLeuAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIle 2306

Db 542 AAGCAAGAGATCATCAAGAGCTGCACATCAACCAACCGATGAGCCCTGATACATATC 601

Qy 2307 SerGln 2308

Db 602 AGCCAA 607

RESULT 100

LOCUS BY763864

DEFINITION BY763864 RIKEN full-length enriched, NOD-derived CD11c +ve dendritic cells Mus musculus cDNA clone F630016H22 3', mRNA sequence.

ACCESSION

VERSION BY763864.1 GI:27200956

KEYWORDS

SOURCE EST.

ORGANISM Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 676)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojibori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuoka, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,

Chothia, C., Corbani, D.E., Cousins, S., Dalla, E., Dragani, T.A.,

Flatcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maita, L., Marchionni, D., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Hirozane, T., Hori, F.,
Fukuda, S., Hashizume, W., Hayashida, K., Kagawa, I., Kawai, J., Kojima, Y.,
Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES

SOURCE

Location/Qualifiers
1..676
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="F630016H22"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched, NOD-derived CD11c
+ve dendritic cells"

ORIGIN

Alignment Scores:
Pred. No.: 9,03e-33 Length: 676
Score: 1011.50 Matches: 191
Percent Similarity: 90.62% Conservative: 12
Best Local Similarity: 85.27% Mismatches: 18
Query Match: 7.65% Indels: 3
DB: 13 Gaps: 1

US-09-522-753-5 (1-2517) x BY763864 (1-676)

Qy 1387 ProProSerArgAspLeuThrGluAlaTyrLysThrGln-----AlaLeuGly 1403

Db	4	CCACACACCTNAGGACCTGACTGACANTACAAGCNCGGGCCCTTGGACCCCTCTGGGT	63
Qy	1404	ProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAlaGlyArg	1423
Db	64	CCCTGAAGCTGAAGCCGACTCACGAGGGTGTGGTAGCAACTGTGAAGAAGCGGGCCG	123
Qy	1424	SerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuProLeuAlaPro	1443
Db	124	TCTATCCATGAGATCCCGAGAGAGAGCTGCGCGCACACCTGAGCTACCCCTGGCACCA	183
Qy	1444	ArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAla	1463
Db	184	CGGCGCTCTGAAGGAGGGTTCATCACCCAGGGCACCCCACTCAAGTAGCACTCTGGGCA	243
Qy	1464	SerThrGlySerLysLysHisAspValArgSerLeuIleGlySerProGlyArgThr	1483
Db	244	CCCTCCACTGGCACCAGAAACAGCAGTGGGTCAATCATCGCAGCAGCCCGCGCGCCT	303
Qy	1484	PheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGluArgAlaCys	1503
Db	304	TTCCCTGCCCTGCACCCCGCTGGACATAATGGCTGACGCCCGGCACCTGGAGCGTGCCTGC	363
Qy	1504	TyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAla	1523
Db	364	TATGAAGAGAGTCTGAAGAGCCGGTCAAGGACCAGCAGTAGTGGTGCAGGGGGCTCCATCAC	423
Qy	1524	ArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyr	1543
Db	424	CGTGGGGCTCCAGTCTGTCGTGCTGACTGGCAGCCAGCCAGCCCACTGACTTAC	483
Qy	1544	GluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProValThrMetArg	1563
Db	484	GAAGACACAGGGGACCGCTTCCACAGTCACCTGCACGTGGCTCCCTGTGACACGAGG	543
Qy	1564	GluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArg	1583
Db	544	GAGCCCCACGCCCTTCAGGAAGGAGCGCTCTTCACAGCAAGGGCGTCCCAAGGACCGG	603
Qy	1584	LysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrValProGluHis	1603
Db	604	AAGCTGACATCTACACCCCGGAGATCGCCAAAGTCCCCACACAGCACTGTGCCGAGCAC	663
Qy	1604	HisProHisPro	1607
Db	664	CACCTTCACCCC	675

Search completed: April 16, 2004, 08:14:39
Job time : 12777 secs